A NETWORK AND SECURITY ANALYSIS OF THE U.S. INTERNET BACKBONE NETWORK

By

ANGELA MCINTEE

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2004

TABLE OF CONTENTS

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Doctor of Philosophy

A NETWORK AND SECURITY ANALYSIS OF THE
U.S. INTERNET BACKBONE NETWORK

By

Angela McIntee

May 2004

Chairperson: Timothy J. Fik
Major Department: Geography

Currently, as both the public and private sectors have become increasingly reliant on Internet-related infrastructure, it is essential that the most valuable components of the telecommunications system be identified and protected from disruptions and sabotage, to ensure the proper functioning of the nation's economy and communications networks. Any disruption that might lead to the loss of a network component could have devastating consequences for both the overall network and the economy at large. In light of these concerns, our study focuses on U.S. telecommunication networks and the Internet backbone. Our study highlights the importance of spatial variability and discusses the potential susceptibility of network components. More specifically, our study outlines a methodology for identifying critical nodes and links in the U.S. Internet backbone network. This type of analysis will aid policy-makers in the allocation of resources when determining which infrastructures are most important to protect or duplicate, to minimize the threat of a disruption. Identifying critical components is essential for prioritization schemes that may be developed to add redundancy and

circuits to the network, to ensure its proper functioning in the event of an attack or disaster. A graph theoretic approach is used to define and rank nodes and links and to measure their importance to the overall network using both weighted and unweighted scenarios. Implications of various node- and linkage-removal scenarios are also discussed. Empirical results suggests clustering of telecommunication infrastructure and bandwidth within large metropolitan locations, with regional variations in connectivity that are not simply a matter of population size Understanding the Internet as a "network of networks" will aid in protecting and preserving the network, lessen component susceptibility to disruptions, and enhance its overall efficiency.

CHAPTER 1
PUTTING THINGS INTO PERSPECTIVE

Introduction

On January 25, 2003, the Sapphire worm was launched by malicious hackers, and reached global diffusion in 10 minutes. In the process, several bank automated teller machine (ATM) networks failed, airline reservation systems ground to a halt, and the entire Internet suffered a global slow-down. Results of the Sapphire worm show not only the interdependency of telecommunication networks, but remind us of their vulnerability. Before this cyber-attack, events of September 11, 2001, brought a heightened sense of awareness and increased focus on the security of critical infrastructure in the U.S. As the world becomes increasingly connected, because of the Internet and advanced telecommunications technology, more attention has turned to issues of cyber terrorism. The increasing dependencies on the ever-expanding telecommunications and information technology (IT) have brought new concerns over security and susceptibility to attacks. In the current atmosphere of tense international relations, it is now imperative that law, policy, regulation, and technology are more fully integrated in the field of telecommunications This will help to ensure the stability and security of the nation's critical, sophisticated, and valuable information infrastructures. Understanding the geography of this infrastructure is imperative: the nation's economic security is highly reliant upon this vital resource to support expanding financial networks and information needs.

This research explores the location and connective properties of the Internet and advanced telecommunications networks, concentrating specifically on physical infrastructure, and more specifically, on the Internet backbone network. Though the

current role of urban planners and local government is minimal in the building of telecommunication infrastructure, decision makers in the telecommunications industry, as well as urban policymakers at the national level, have recognized the importance and need for policy and regulation.

Today's Internet is comprised of various infrastructures. The impact of telecommunications infrastructure and the Internet on urban systems, businesses, academia, government, and consumers has increased dramatically since it's commercialization. Users have become more reliant on Internet infrastructure and technology to carry out basic functions and communication activities. Subsequently, they have become more vulnerable to network disruption. The first step in answering questions of vulnerability and risk related to telecommunication networks is to incorporate different types of telecommunication data into one common information system.

This research project includes various types of telecommunication and Internet infrastructure data. The research also looks at present policy governing telecommunication networks and infrastructure, as well as pending policy that will further impact the industry. To aid in our understanding of this problem, it is imperative to determine the geographical locations of those links and nodes that are most valuable to the telecommunication industry- links and nodes that would have the greatest negative impact should information flows through these assets become disrupted. The results of the analysis will then be placed in the context of national and regional security. Identifying the physical location and interdependencies of critical links and nodes will allow for security emphasis to be redirected to those network components that are currently most vulnerable to an attack, thereby creating a prioritization scheme for creating greater redundancy to potentially minimize the impact of node or linkage disruption. The methodology employed in this dissertation may be widely adopted time and time again to assess this network.

Statement of the Problem and Hypotheses

This research was intended to explore telecommunication network connectivity and vulnerability. Currently, one of the most pressing issues of homeland security in the U.S. is the protection of critical infrastructure. In response to growth concerns over domestic terrorism, this research focuses on the telecommunication network and geographic variations in its susceptibility to an attack. Extra insurance in the form of protection and prevention is required to preserve the most-valuable links and nodes of this network. And hence, the locational aspects of the most valuable components must be identified. Any disruption that might lead to the loss of a node or link in a network could lead to devastating consequences upon the overall network. Determining the most critical links and nodes will allow for increased protection of the infrastructure.

Preliminary research on this project indicates that highly connected nodes will be the most critical to the overall network. It is hypothesized that the most important nodes will house the most bandwidth. It is also hypothesized that the most highly connected nodes presently house multiple interconnection facilities called colocation facilities (a more detailed description of these facilities can be found in Chapter 2); the demand for interconnection increases positively in direct proportion to fiber bandwidth. Although these nodes may be the most directly connected in the network, they are not necessarily at the top of the nodal ranking in terms of both direct and indirect connectivity. It is further hypothesized that the most critical links will be connected to the most critical nodes.

This dissertation also will identify critical clusters of Internet activity on the east and west coast of the U.S. and other regional subnetworks or clusters. It is hypothesized that the most highly connected places contain the largest amount of bandwidth and highest number of connections to other places, thus housing the most infrastructure. There is, however, a definite variability in the prominence of links and nodes, as well as a varying

degree of vulnerability. The complexity of a network multiplies as the number of links and nodes comprising the network increases. As the complexity and size of the network may vary, so might the vulnerability of each link and node within the overall network. Links and nodes that are more critical to the overall network will have a greater impact on the network should they become disconnected from the network. Furthermore, it is hypothesized that the overall impact of the removal of a node or linkage may not be obvious without a more in-depth analysis that considers all direct and indirect connections. Connectivity indices will be used to determine the most critical links and nodes in the network. It is hypothesized that places that are less prominent to the overall network, but more highly connected in a regional sense may be more vulnerable to an attack and potentially more disruptive.

And finally, it is hypothesized that there will be great variability in the overall effect of disruption or termination of flows through various nodes and linkages with regional implications that are not obvious. When a link or node is damaged or disrupted, the connective properties of nodes and links may change completely. By applying graph theory to analyze the network, the structure and properties of the network will be highlighted and components tested. This will allow for the identification of circuits and redundancy that would increase the overall connectivity of the network and minimize the impact of disruption.

The Internet

History of the Internet

The Advanced Research Project Agency network (ARPANET) was created in the late 1960s as a network project of the Advanced Research Projects Administration (ARPA) under the direction of the Defense Department. The Internet is a byproduct of that project, and though appearing as a recent phenomenon, it is actually a representation of decades of development (Abbate 1999). By the mid-1980s the "Net,"

as ARPANET was nicknamed, had transitioned into the hands of the National Science Foundation (NSF). The NSF allowed the Net to be used strictly for academic and research purposes (Boardwatch 1999). However, private firms and corporate developers of early Internet technology such as: Bolt, Baranek, and Newman (BBN)—a military contractor and consulting firm involved in the early engineering and development of the project—realized the economic potential of this research network (Schiller 1999). Firms developed networks and collaborated with one another by interconnecting these new networks to create a private-sector version of NSFNET for their corporate clients, essentially duplicating the NSF's Net. The Internet switched from an academically oriented network to a commercially oriented one in 1991; when the NSF decided to allow commercial traffic across NSFNET (Thomas & Wyatt 1999).

Today the Internet is heavily used for e-commerce: shopping, trading stocks, pornography, music downloading and file sharing and real estate, as well as a valuable information resource tool. The Internet has become an important forum for news media Major broadcasting companies have websites with audio and video clips updated around the clock, and newspapers have many sections of their printed version posted on websites. The Internet has also radically enhanced personal communication via: e-mail, chat rooms,[1] instant messaging.[2] The Internet has evolved from a single experimental network serving a dozen sites in the United States to a "network of networks" linking millions of computers and servers worldwide (Abbate 1999).

---

[1]A chatroom is a place or page in a Web site or online service where people can "chat" with each other by typing messages that are displayed almost instantly on the screens of others who are in the "chat room." Chat rooms are also called "online forums."

[2]Instant messaging is a service that alerts users when friends or colleagues are on line and allows them to communicate with each other in real time through private online chat areas.

Infrastructure and the Net

The network of networks is comprised of links and nodes, a global network connecting millions of computers worldwide to exchange data and news. There are several levels of networks at work. Interconnection of these networks is key to the functionality of the overall network. The Internet backbones are the long-haul routes that link the nodes of the Internet to users.

The Core of the Global Net Is Centered within the U.s.

Although telecommunications began in America with the introduction and diffusion Samuel Morse's telegraph, the U.S. has not always been the world leader of telecommunications networks. Great Britain developed superior telecommunication technologies and communication networks early on in the 1870s (Hugill 1999). Abler (1991) attributes the development of America's telecommunications industry to system-specific software and a series of historic accidents, such as the failure of the U.S. Congress in 1845 to see any value in its ownership of the patent on the telegraph. During the formative stages of the Internet's development, the United States established itself as a world leader in implementing the technologies and infrastructure needed to develop and nurture this innovation. This would explain why, on a global scale, the U.S. became the center of Internet activity (Cukier 1998, Finnie 1998, Malecki & Gorman 2001).

Cukier (1998) gives several reasons why the Internet is "U.S.-centric": (1) It had a "head-start in building infrastructure and guiding the location of Internet content; (2) the artificially high cost of cross-border capacity outside the U.S.; and (3) and customer demand for Internet service" (p. 113). Dodge and Kitchin (2001) describe the distribution of Internet users, stating that in February of 2000, the U.S. and Canada accounted for 136.06 million users with approximately 5% of the world's population; while Asia and Europe *combined* accounted for 126.89 million users despite accounting for well over

60% of the world's population. Internet traffic patterns indicate that over half of European Internet traffic and 70% of Asian Internet traffic travels through the United States. In 1999, the United States housed 58% of Internet hosts and content, and only 6% of the 100 most visited web sites were located outside of the U.S. (Cukier 1998). Although the U.S. has a significant lead in telecommunications infrastructure, it is likely European telecommunication growth will soon boost Europe to rival the U.S.

Major metro areas in the U.S. and the telecommunication companies that provide Internet infrastructure within them are in a fierce competition for top ranking. Based on Internet activity and infrastructure, these cities are in a constant shift of rank, depending on what type of infrastructure is being measured. However, the same seven metro areas continuously make the list: namely, New York, Washington, D.C., San Francisco, Chicago, Dallas, Los Angeles, and Atlanta (Cukier 1998, Finnie 1998, Graham 1999, Malecki & Gorman 2001, Moss & Townsend 2000). New York is the leading Internet hub in the global economy, housing nine fiber networks, more than any other metropolitan area (Finnie 1998).

The metropolitan rankings compiled by Atkinson and Gottlieb (2001) do not match rankings that measure strictly telecommunication infrastructure because their research includes other economic variables. The *Metropolitan New Economy Index* (Atkinson & Gottlieb 2001) has compiled a ranking of U.S. metro areas based on five subcategories: knowledge jobs, degree of globalization, economic dynamism and competition, the transformation to a digital economy (including infrastructure), and technological innovation capacity.

Internet Links—Backbones and ISPs

The term "Internet service provider" (ISP) is an overgeneralization that combines both small, local Internet service providers and globe-spanning Internet backbones, and the term does not differentiate among the various types of ISPs. An ISP is a company

that provides access to the Internet, serving individuals as well as large companies, with direct links. ISPs can play vastly different roles. ISPs can be part of a major backbone and a global network, or they may be local providers leasing infrastructure and servicing a limited geographic market.

An Internet "backbone" simply can be defined as a collection of wires that connect the Internet's nodes, linking them together so that they may exchange data. In more complex terms, the Internet backbone is defined by the National Telecommunications and Information Administration (NTIA) as "a set of paths that local area networks (LANs) connect for long-distance connection. A backbone employs the highest-speed transmission paths in the network. A backbone can span a large geographic area. The connection points are known as network nodes or telecommunication data switching exchanges (DSEs)" (Telegeography 2001, p. 102). International backbones were defined by Telegeography (2001, p. 102) as "Private data links which cross international political borders, run the Internet Protocol (IP), are reachable from other parts of the Internet and carry general Internet traffic: e-mail web pages, and most of the other popular services which have come to define today's Internet" (Telegeography 2001, p. 102). A backbone firm, such as Sprint, may serve as an ISP itself, or an ISP may lease access to the Internet from a backbone provider. At the same time, many different ISPs utilize the same fiber. This means that different communications links, even when obtained from different providers, many run over the same fiber, in the same bundle, or in the same conduit (NRC 2001).

Cukier (1998) classifies Internet Service Providers (ISP) into four groups: (1) backbone ISPs; (2) downstream ISPs; (3) online service providers (e.g., American Online, Microsoft Network); and (4) firms specializing in Website hosting (e.g., Qwest) (Table 2-1). Backbone ISPs include those that may connect the Internet globally and

transfer the largest amounts of data, such as Exodus, Globix, Sprint, MCI WorldCom, and IBM.

Table 2-1. Internet service provider groups

ISP Group	Example
Level 1: National backbone ISPs	SprintLink, MCI
Level 2: Downstream ISP	AOL TimeWarner, WorldCom
Level 3: Online Service Provider	America Online, Microsoft Network (MSN)
Level 4: Website Hosting	Qwest

According to Malecki and Gorman (2001), 48 national backbone operators currently operate the transit network in the U.S. The next level is comprised of the downstream ISPs, hundreds of local and regional ISPs that serve mainly individuals and small and medium businesses. Malecki & Gorman (2001, in Brunn & Leinbach 1991, p 91) divided the Internet hierarchy into five levels (Table 2-2). All networks exchange data on the first level. The second level makes the transfer of data possible among cities around the world. Regional networks comprise the third level, though Malecki and Gorman warn, "they [regional network providers] may be a dying breed as they are replaced by national providers" (Malecki & Gorman 2001, p. 92). Internet Service Providers (ISPs) are the fourth level. Finally, Internet users are the fifth level.

Table 2-2. The hierarchy of Internet network interconnections

Level		Providers	Example
Level 1:	Interconnection	Network Access Points (NAPs), private peering points	Ameritech Chicago NAP
Level 2:	National backbone	National backbone operators	SprintLink, MCI
Level 3:	Regional networks	Regional Network Operators	Erols, Rocky Mountain Internet, Inc. (RMII)
Level 4:	ISPs	Internet Service Providers	DialNet, bright.net
Level 5:	Users	Business and consumer market	

"Middle-Mile" Network Links

These are the links of the "aggregators," firms that connect large data customers, such as firms in office buildings and office parks, to local points of presence (POPs) of backbone networks. Many utilities such as Gainesville Regional Utilities (GRU) and Florida Power and Light (FPL) serve as middle-mile providers in this service area. Middle-mile facilities connect the backbone fiber owned by large national telecommunications firms, such as Sprint or MCI WorldCom, to regional networks that may be owned by utility companies, or smaller telecommunication firms. Middle-mile facilities are an integral part of the Internet's backbone hierarchy, providing linkage between national/regional networks and local networks. A recent report by the Federal Communications Commission (FCC) (2000) includes middle-mile facilities as one of three main types of the Internet's network components (backbone facilities and last-mile[3] facilities comprise the other two). Thus far, little or no research has been done to analyze this scale or segment of the Internet.

The Last-Mile Connection: From Backbone to Computer

The "last-mile" connection refers to the connection between the Internet backbone and user; it is the final physical linkage between user and network. A user's connection to the Internet may be one of three types: dial-up, continuous, or wireless. Continuous connection is the most efficient connection, enabling instant delivery of e-mail, eliminating the need to tie up a phone line, and allowing businesses to advertise and publish directly to the Internet and enable "real-time energy management" (Hurley & Keller 1999, p. 3). Most households currently access the Internet through dial-up connections, using a modem and a telephone line. This method is the cheapest way to access the Internet. Users may not have the option to use a faster, more expensive type

[3] "Last mile" is the term used to describe the connection between the user and the ISP, which is typically the slowest aspect of Internet access.

of connection because access to sophisticated types of Internet infrastructure are not offered in their neighborhood. Integrated Services Digital Network (ISDN), developed in the early 1980s to improve telephone service, is a technology introduced in the early 1990s that provides moderate bandwidth (64 to 128 Kbps[4]). This technology was efficient for connecting to the Internet, though a slow and lengthy process. Soon after, more technologies emerged. cable modems offering high-bandwidth and continuous connection (10 Mbps[5] or 30 Mbps), Asymmetric Digital Subscriber Line (ADSL), a high-bandwidth copper-wire technology (8 192 Mbps/640 Kbps up). These new technologies boasted higher bandwidth, which allowed for a higher data transfer rate and volume. The connections were dedicated to data transfer, allowing the user to maintain a constant, uninterrupted transfer.

Physical distance matters with many types of sophisticated telecommunications infrastructure, such as digital subscriber line (DSL). The closer one is to the service provider, the better the service (Moss 1998). With a full range of options for providing high-bandwidth local access, it is clear that no single technology will be declared the "all-around winner" (Hurley & Keller 1999, p. 37). High-bandwidth technologies remain competitive, with user preference between cable modems and DSL not apparent. Wireless is the newest form of access to the Internet. The development and deployment of wireless service to provide mobile access hinges on the results of current FCC efforts to open up radio-frequency spectrum for such services (NRC 2001). Currently, a wireless connection to the Internet is not considered secure and may be decrypted by hackers quite easily. This lack of security greatly deters wireless users from making financial transactions or transferring valuable data via a wireless connection. Whichever

[4] Kbps- kilobits per second (thousands of bits per second). Kbps is a measure of bandwidth (the amount of data that can flow in a given time) on a data transmission medium.

[5] Mbps-megabits per second (millions of bits per second). Mbps is a measure of bandwidth on a transmission medium.

technology is used to connect to the Internet, the user is able to navigate the Internet, check e-mail, and communicate with other users. As data are requested, users may navigate from one network to the next to reach the data source, call up the data packets, and navigate back through the Internet to complete the request.

Internet Nodes

The original nodes of the Internet were universities and research institutions invited by the Department of Defense to participate in the ARPANET project, they included prestigious research institutions such as MIT, Carnegie Mellon, and UCLA, to name a few. These universities were given large grants in 1965 to create "centers of excellence" computing research centers. These research centers, dispersed throughout the United States, were connected to form ARPANET. It was the original users that transformed ARPANET into the Internet we know today. The users of the network could create new applications with few restrictions and had the incentive and ability to experiment with the Internet to mold it to better meet their immediate needs, for example, building new hardware or software, or using the existing infrastructure in new, improved ways (Abbate 1999).

Large businesses, universities, and larger institutions often have direct links that allow the user to bypass the telephone network and connect directly to a metropolitan-based network or to Internet backbones (Malecki & Gorman 2001). Langdale (1989) explains that large global companies that lease sophisticated networks that utilize high-speed circuits dominate international business telecommunications traffic, allowing the firms to link their networks to networks housed in major industrialized countries. The cost of communication networks is largely determined by the maximal capacities of networks, but the traffic those networks carry depends on how heavily those networks are used. Thus, increasing the efficiency of data transport would make the Internet less expensive and more useful (Odlyzko 2000).

The Internet today is an amazing network, but the individuals, organizations, government, businesses, and educational institutions incorporated within and linked by the Internet make it the invaluable resource it has become The network is a medium for the exchange of information, data, and ideas among those who use it. The Internet is the most influential advancement in the distribution and exchange of information since the telephone (Moss & Townsend 2000).

Network Interconnection

In February 1994, the National Science Foundation (NSF) designated four nodes as Network Access Points (NAPs): San Francisco, operated by PacBell; Chicago, operated by Bellcore and Ameritech, New York, operated by SprintLink (this NAP is actually in Pennsauken, New Jersey); and Washington, D.C., operated by Metropolitan Fiber Systems. NAPs are "sites where private commercial backbone operators could interconnect" (Boardwatch 1999, p. 13). When NSF completely transferred responsibility and rights to the Internet to commercial entities in 1995, the networks interconnected only at the NAPs. When the NSFNET backbone was shut down and transferred to the commercial entities, the NAP architecture became the Internet (Boardwatch 1999). Each of these four NAPs would be maintained and operated by telecommunication companies rather than the NSF.

As the Internet continued to grow and develop, traffic increased dramatically, and the NAPs became increasingly congested and utilized; demand for more interconnection points began to increase. The core of U.S. Internet interconnection remains the four "official" NAPs and includes other major connection points, the Metropolitan Area Exchanges (MAEs), Boardwatch 1999, p. 13). Thirty-eight of 41 major backbone networks in the U.S. connect at both MAE East and MAE West (Malecki 2000). From this, we can conclude that the importance of the NAPs has not declined. The MAEs and NAPs were and are considered public facilities where backbones and ISPs could

interconnect and colocate at little or no cost. However, because a large number of networks link at each NAP, congestion and inefficiency are common.

The solution to the demand for more efficient, faster connections was private interconnection; a term originally coined "private peering." Network peers connected at private locations rather than at the NAPs. The term private peering is sometimes misused, describing network interconnections in general, whether the networks are equals or gross unequals (for example, national Internet backbone provider networks would not be equal to a local Internet service provider). Private interconnection is a relationship between two or more ISPs in which the ISPs create a direct link between each other and agree to forward each other's packets directly across this link instead of using the standard NAPs, or Internet backbone. Simply put, peering takes place between network equals, and interconnection takes place between unequal networks, with the weaker party paying for transit. Interconnection can involve more than two ISPs. In this situation, all traffic destined for any of the ISPs is first routed to a central exchange, which is called a peering point, and forwarded on to the final destination after hitting the peering point. Private peering points function similar to the larger interconnection points because they provide interconnection between networks. Colocation facilities operate on a smaller scale than the NAPs, IX facilities, and MAEs with contracts and higher fees for the users (Boardwatch 2000). The commercial Internet operates as a machine of collaboration and cooperation between networks. "Every ISP network must inter-operate with neighboring Internet networks in order to produce a delivered outcome of comprehensive connectivity and end-to-end service" (Huston 1999, p. 1).

This chapter has discussed the relevance of this dissertation in relation to current concerns of the protection of critical infrastructure, particularly telecommunication networks. A large disruption to the Internet will cause large repercussions as it is directly

tied to the economy This dissertation determines the most important links and nodes to the Internet backbone network by employing network analysis. It is hypothesized that the links and nodes most critical to the network will have a high concentration of fiber bandwidth and an equally high concentration of colocation and interconnection facilities.

Chapter 2 presents a review of literature relating to networks, geography and telecommunications. Chapter 3 explains the data, analysis, and methodology used in this research. Chapter 4 presents and discusses an unweighted analysis of the U S. Internet backbone network Chapter 5 presents and discusses a weighted analysis of the U.S. Internet backbone network. Chapter 6 presents a statistical analysis of network measures, using indices derived from Chapters 4 and 5. Chapter 6 also summarizes the findings and discusses future research possibilities of the U.S. Internet backbone network.

CHAPTER 2
NETWORK CONCEPTS: A LITERATURE REVIEW

In order to fully understand the research implications and to identify related research to further aid in answering the proposed questions, this chapter examines relevant literature and reviews basic concepts in network analysis. The initial portion of this chapter will focus on network concepts and constructs, with a brief overview of networks. The second section reviews geographic literature of networks and information, concentrating on network analysis, telecommunication infrastructure analysis and the Internet backbone. The third section has two subsections: historical telecommunications and the Internet, highlighting the importance of telecommunications to the city.

Network Concepts and Constructs

A short glossary of network terminology and concepts is provided below to facilitate the discussions and analysis. These definitions will serve as a point of reference and to help clarify various network concepts. A network can be defined most simply as two or more nodes connected by link(s), an interconnected system of objects or people, and consist of a set of links and nodes. Nodes and links are the main composition of a network: nodes/vertices are connecting points or objects in a graph or network while links/edges are the connections between them.

Graphs are abstract representations or models of a network. The terms vertices and edges are most commonly used when describing a graph, a vertex refers to a node and an edge refers to a link. In this text, vertices and nodes, and links and edges will be used interchangeably in this text. Relative graph theory applies abstract configurations that consist of points and lines to study network properties. A directed graph has ordered

pairs of edges connected by links with direction while an undirected graph has unordered pairs of edges connected by links without direction.

A network component refers to which set of nodes a vertex belongs that can be reached from it by paths running along edges of the graph. A geodesic path is the shortest path through the network from one node to another, while the diameter indicates the longest geodesic path between any two nodes (Kochen 1989). A circuit indicates the flow of a path. Circuits denote direction of a path or route within a network, that lead back to themselves.

A network connection or the connectivity of a node describes the relationship between nodes or links in a network, a topological description that specifies the interconnections between nodes. Connections may be direct or indirect. Accessibility is a description that describes the degree to which a node can be reached or accessed by other nodes in the network, given their absolute or relative locations Location-based accessibility is the degree to which a node can be reached or accessed given its location in relation to the other network nodes, as based on physical distance and spatial structure.

The nodality of a node refers to the degree of a node's dominance within a network. To measure the location of a network's components is to determine the centrality of a node within the network (Kochen 1989, Perrucci & Potter 1989). A gateway is a nodal characteristic that describes a main entry-and-exit point for a region or network. A disconnected node or segment has been detached or removed from the network or subnetwork

Structural equivalence is a measure of the similarity in roles of nodes in a network, through the determination of which nodes play similar roles in the network. For example, in Figure 2-1, nodes V1 and V3 are structurally equivalent, playing similar roles within their regional cluster of nodes as collector nodes, or regional hubs. Structural

holes are areas of no connection between nodes that could be used for advantage or opportunity.

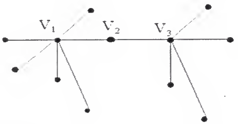


Figure 2-1. Network A

Network centralization is equivalent to a node's location or position in a network considering all direct and indirect links, or from a multi-faceted standpoint, how a node is characterized in terms of connectivity, accessibility, it's propensity to cause disconnects upon it's removal or failure, and its importance in terms of adding redundancy or circuits as a back-up node should other nodes (nearby) become removed or fail.

"Betweenness" is a measure of influence over what flows in the network, a measure of power that a node has based on its relative location or position in the network. For example, in Figure 2-2, node V3 is an important node in that it helps connect two distinct regional clusters of lesser-connected nodes. Note that the removal of node V3 would cause the network to become disconnected (Krebs 2004).

"Closeness" is a measure of how far any node is to any other node in the network. For example, Nodes V3, V7, and V8, in Figure 2-2, are only two links away from any other node in the network (max). Tier 1 in the nodal hierarchy in Figure 2-2 would be comprised of nodes V3, V7, and V8. The second tier would consist of the remaining nodes, V1, V2, V4, V5, V6, and V9. Note: Closeness is different from diameter in that diameter (d=3 for this network) representing the minimum number of links between the two most distant points in a network (Krebs 2004).

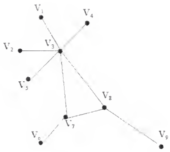


Figure 2-2. Network B

"Boundary Spanners" are nodes that are more central than their immediate neighbors whose connections are only local. In short, they are regional hubs or the center or predominant node in a regional cluster or subnetwork (Krebs 2004).

"Peripheral Players" are nodes that are often connected to outside or external networks that are not currently mapped, making them of grater importance than nodes that are not directly linked to those external networks. Hence, their importance may be understated (Krebs 2004, Kochen 1989).

Network Categorization

Networks may vary in nature, size, and purpose. They may range from the environment, to human, to technical, from tangible to intangible and can be physical or virtual. Networks might be human, technical, or natural, and private or public. Several disciplines use study and utilize networks for various purposes; transportation, landscape ecology, geography, neurology, telecommunications, communications, physics, computer science, economics, health care and medicine, electric and gas, water distribution and resources, urban planning, mining and geosciences. Networks can be organized into various categories. The following section discusses physical and virtual networks, technical networks. environmental and natural networks, social and human networks, and private and public networks

Physical and virtual networks

Physical networks are two or more nodes that are connected by a physical link; roads, wires, corridors, streams, cables, pipes. The U.S. Interstate Highway System is an example of a physical network.

Virtual networks exist conceptually, rather than being physically real. Though a virtual network may not have a physical composition, it still serves the same purpose as physical networks; to connect various objects for the sharing of data and knowledge. Networks can be both physical and virtual in nature, the prime example being the Internet (www.webopedia.com). Some social networks might also be considered both physical and virtual; meeting physically or being virtually connected by association.

Technical networks

Technical networks are used in science or industry and are of a mechanical nature. They typically include sophisticated equipment that may be electronic or computerized. An electric company providing power to a neighborhood is an example of a technical network. The power-grid itself is public, using the power plant, high-voltage transmission lines, power substations, transformers, power poles, and transformer drums to move power across a network to consumers. All of the equipment is electrical, mechanical, or computerized in nature.

Optical networks are high-capacity telecommunications networks based on optical technologies and components that provide routing, grooming, and restoration at the wavelength level as well as wavelength-based services. Optical networks are providing more advanced capabilities as well as lower costs for the telecommunications industry. "In the early 1980s, a revolution in telecommunications networks began that was spawned by the use of a relatively unassuming technology, fiber-optic cable. Since then, the tremendous cost savings and increased network quality has led to many advances in

the technologies required for optical networks, the benefits of which are only beginning to be realized" (IEC 2002).

Natural and environmental networks

Natural networks include environmental networks. As environment encompasses disease and epidemiology, so percolation of disease on a network falls in this category. The transfer of biological agents occurs through natural networks. With the sudden onset of Severe Acute Respiratory Syndrome (SARS) in early 2003, we have been reminded of the dangerous ability of disease to travel rapidly through a network of individuals or entities. The Center for Disease Control (CDC) advised against unnecessary travel to those areas infected with SARS (CDC 2003), in an attempt to prevent the transmission of the disease through transportation and social networks. Other natural networks include hydrological networks. Hydrological networks describe the movement and interaction of water, including lakes, streams, rivers with each other and other environmental impacts.

Neural networks are comprised of processing elements called units that respond in parallel to a set of input signals given to each. Neural networks are most closely associated with the study and research of the human brain as the unit closely represents the brain's neuron. In the medical field, neural networks are called referred to as artificial neural networks (ANNs) and have been in existence since the 50s and are used in a wide variety of applications, including speech recognition, which was the original intent of creating an ANN "Artificial neural networks (ANN) are a very simple model of the brain" (O'Sullivan & Unwin 2003, p. 364). ANNs are now used in geographical applications. ANNs have geographic applications. They are based on the idea that "brainlike structure = intelligence." This type of network assumes that it can operate in two types of environments: supervised or unsupervised. A supervised network is a network that has been trained or programmed by a known set of data. An unsupervised

network operates in a more traditional form, in that it eventually "settles to a state such that different combinations of input data produce different output combinations that are similar to a clustering analysis solution" (O'Sullivan & Unwin 2003, p. 365).

Social and human networks

Social networks are formed as a means to communicate with and within a group. The nodes in social networks represent people or groups, while the links represent relationships or flows between them. Examples of social networks might include a school, political organization, a firm, unionized workers, and friends. Other types of human networks may not be as obvious as others. For example, transportation networks could be considered human. Transportation networks are built by for the purpose of transferring goods and information. However, in order to create one network, the disruption of another may occur. Road and railroad networks may cause landscape fragmentation that might affect other types of networks. Recent research has focused on the environmental impact of road networks, and further the impact upon animal behavior and population (Carleton 2003). A review of relevant social network analysis follows later in this section, within the review of complex network topology literature.

Private and public networks

Public networks can be used free of cost while a private network charges some types of user fee. Government agencies, business and individuals can, and do, set up their own private networks. Private networks allow for security and privacy by controlling user access to the network. Private networks generally require a user fee. There are a number of public networks, both large and small that are commonly available to government, business and consumers in the USA. Some of the most common public networks are: The Public Switched Telephone Network (PSTN, wired telephones) Public Wireless Voice Networks (PWVN, such as cellular and Personal Communications Service [PCS]) Public Wireless Radio Paging Networks (PWRPN), and the Internet.

Private networks, such as government and corporate, may or may not connect to public networks, such as the Internet. However, these public and private networks increasingly overlap. Private and public networks may be physical, virtual, or both. A private virtual telecommunication network allows for an individual or entity to remotely access a larger network using public network infrastructure, while maintaining privacy and security by encrypting data that is exchanged.

The "Small World" Phenomenon

Many have studied the small world phenomenon including, but certainly not limited to Milgram (1967), Kochen (1989), Wasserman and Faust (1994), Ozana (2001), Albert and Barabasi (2002a), and Watts (2003). The small world concept is that, though a network may seem vast in size, the nodes in the network are connected with short paths. Milgram (1967) pioneered the concept and the most popular theme "six degrees of separation." Milgram concluded that there was an average of six paths between most pairs of people in the U.S.

Network Analysis

Complex Network Topology Literature

This group of network literature discusses the exploration of the topological properties of real networks. Complex network analysis is used in a variety of disciplines to understand various networks and real systems (Albert & Barabasi 2002b, p. 49). Albert and Barabasi (2002b) give a very detailed review of complex network research in their paper titled *Statistical Mechanics of Complex Networks*, published in the Reviews of Modern Physics. Newman (2003) also gives an excellent review of complex network analysis in his publication *the Structure and Function of Complex Networks*. These two papers were used as a guide for the following review of examples complex network research.

Social networks

The social network serves as a major source of social capital. Being part of a social network means to communicate with and within a group. Examples of social networks might include a school, political organization, a firm, unionized workers, and friends. Two techniques for approaching social networks are described by Watts (2003, p. 48): network structure and social structure. Sociologist Mark Granovetter (Granovetter as cited by Watts 2003, p. 49) concluded that effective social coordination does not emerge from strong ties to a social network, but rather from occasional weak ties. Granovetter described in his 1973 paper, The Strength of Weak Ties, a foresight of what is now described by Watts (2003, pp. 49-50) as the new science of networks. Watts concludes that "social network analysis still has one major glitch; there is no dynamics" (Watts 2003, p. 50). Thus the measure and study of social networks is complex and problems are approached differently, depending on the nature of the network and it's components.

Actor collaboration network

One of the most analyzed social networks is the movie actor collaboration network. This network contains all movies and the casts of these movies since the 1890s. The network is continuously expanding and updated and is based on the Internet Movie database. Watts and Strogatz (1998); Newman, Strogatz and Watts (2001); Barabasi and Albert (1999); and Albert and Barabasi (2000) have all used the movie actor database. Watts and Strogatz (1998) reported that in 1998 the network had 225, 226 nodes (actors). By May of 2000, the number of nodes (actors) had grown to 449, 913 according to Newman, Strogatz, and Watts (2001). When two actors work together in a film, they have a common link.

Science collaboration network

The science collaboration network is similar to the movie-actor network. When two scientists work together, they are connected nodes. Newman (2001a, 2001b, 2001c)

studied four databases during a five-year time frame that included physics, biomedical research, high-energy physic, and computer science to determine the topology of this network. Each of the networks shows a small average path with high clustering coefficients (Albert & Barabasi 2002b). The collaboration network of mathematicians and neuroscientists that published between 1991 and 1998 was studied by Barabasi et al. (2001); they were found to have consistent degree distributions with other collaboration networks.

Sexual contact network

Liljeros et al. (2001) have investigated sexually transmitted diseases (STDs), including AIDS. They studied a network based on the sexual relationships of 2810 individuals. The data was obtained from a Swedish survey conducted in 1996. The distribution of sexual partners was studied for a year. The spread of STDs through the network was studied (Albert & Barabasi 2002b). Due to the fact that the average edge in the network has a relatively short-span, they analyzed the distribution of partners over a one-year period.

Cellular networks

The metabolism of 43 organisms was studied by Jeong et al. (2000). In this project, networks in which the nodes were substrates and the links were chemical reactions represented the organisms. The average path was found to be roughly the same in each of the organisms. Wagner and Fell (2000) looked at the clustering coefficient while focusing upon the energy and biosynthesis metabolism of the E-coli bacterium. Their results show an undirected version of this substrate graph (network) has a small average path length with a large clustering coefficient (Albert & Barabasi 2002b). Protein-protein interactions within a cell were also considered in the analysis, as they help to characterize the cell network. The proteins represent nodes that are connected if they bind together.

Citation networks

The nodes in a citation network represent published scientific articles and links represent a reference to that particular scientific article. This network was studied by Redner (1998). The network included 783, 339 papers cataloged by the Institute for Scientific Information and 24, 296 papers that were published in Physical Review D between 1975 and 1994. The network is formed by citation patterns used within the publications, nodes represent published articles and links represent a reference to a previously published article. Following Redner, Vazquez (2001) did a similar study using the citation network. Vazquez extended the study to include outgoing degree distribution and found an exponential tail.

Linguistic networks

Ferrer i Cancho and Sole (2001), Yook, Jeong, and Barabasi (2001b), and Steyvers and Tenenbaum (2001) are amongst those researchers that study the complex networks formed by human language. Steyvers and Tenenbaum's results indicate that languages form networks and dynamics not so different from other networks (Albert & Barabasi 2002b, p. 53). Ferrer i Cancho and Sole (2001) created a network using the English language, based on the British National Corpus. The nodes represented nodes and were lined to each other if they appeared next to each other, or were one word apart from each other in sentence (Albert & Barabasi 2002, p. 53). The network consisted of 440, 902 words. Ferrer i Cancho and Sole (2001) found that the average path length was small, there was a high clustering coefficient, and there was a two-regime power-law degree distribution (Albert & Barabasi 2002b, p. 53). Yook, Jeong, and Barabasi (2001b) used a different network for their study of the linguistic network. For their network, two words were linked if they were synonyms according to the Merriam-Webster Dictionary. Their results show a large cluster of 22, 311 words out of a total of 23,279.

Ecological networks

Ecologists study food webs or food chains to determine the network relationships between various species. In a food network, nodes represent the species and the links would be the predator-prey relationships between them (Albert & Barabasi 2002b). Williams et al. (2000) recently studied the topology of some of the largest food webs; Skipwith Pond, Little Rock Lake, Bridge Brook Lake, Chesapeake Bay, Ythan Estuary, Coachella Valley, and St. Martin Island. Though the webs were comprised of very different species in different habitats, each indicated that species in habitats are three or fewer links from each other (Williams et al. 2000). The research of Montoya and Sole (2000), and Camacho et al. (2002a) confirmed that the food webs show highly clustered nodes. Montoya and Sole's research focused on Ythan Estuary, Silwood Park, and Little Rock Lake. Two of their research areas overlapped with that of Williams et al. (2000). Camacho et al. (2002a, 2002 b) found that an exponential fit worked well, following the well-documented existence of key species in the food web. They represent a common feature of scale-free networks, hubs Forman and Spearling (2002) explore road ecology. Forman and Spearling discuss the vast network of roads that billions utilize daily. They point out that until now, there has been little or network theory applied to road networks and ecology. The road network and landscape indeed form a complex network. Forman and Spearling did a study of the 4 million miles of public roads in the U.S. and determined how much area they ecologically affect. They concluded that about one fifth of the total U.S. area, 20%, is directly affected ecologically by our road system.

Telephone call networks

The long-distance telephone call network has been studied by Abello, Pardalos, and Resende (1999) and Aiello, Chung, and Lu (2000), amongst others. They constructed a large, directed graph using long-distance telephone call patterns. Phone numbers represent nodes, while every complete call represents a link. These

researchers used the calling network based on the data from one day. They concluded that the degree distributions of the outgoing and incoming edges followed a power law with exponent 2.1.

Power and neural networks

The U.S. power grid consists of generators, transformers and substations, the network nodes The links are the high-voltage transmission lines. With the power outage effecting the northeast U.S. in the summer of 2003, we saw the interconnectedness of this network. The degree distribution of the power grid is consistent with an exponential (Albert & Barabasi 2002b, p. 54). Watts and Stogatz studied the nematode worm, where, the nodes are neurons and a link exists between either a synapse or a gap junction (Albert & Barabasi 2002b, p. 54) In their research, Watts and Strogatz (1998) found that for both networks (power and neural) the average path length was approximately equal to that of a random graph of the same size and average degree, and the clustering coefficient was much higher (Albert & Barabasi 2002b, p. 54).

World Wide Web & the Internet

One of the most recent complex networks to be examined is the Internet. As was introduced earlier, geographers analyze networks, and the geography of networks is often relevant to other disciplines. While geographers were working on early network analysis of transportation networks using graph theory (Kansky 1963, Garrison 1960, Haggett & Chorley 1969), Erdös and Renyi (1960) were focused on theoretical work of complex networks. They modeled large networks using algorithms where N nodes were randomly connected according to probability P. They found that the nodes were connected in a manner that followed a Poisson distribution[1] (Albert & Barabasi 2002B, p. 49). The network model created by Erdös and Renyi was used widely in several

---

[1]The Poison probability distribution is used to analyze how frequently an outcome occurs during a certain time period or across a particular area. Other geographic applications of Poisson involve the analysis of existing frequency count data to determine if a random distribution exists (McGrew & Monroe 2000).

disciplines analyzing networks. The most closely related research of this group would be Internet topology generators (Radoslavov et al. 2000).

According to Barabasi et al. (2001), the absence of topological data in the analysis of complex networks makes random network models the most often applied method of network simulation. As computer technology advanced, and data for real world networks became more available, several empirical finings emerged. Three network characteristics resulted most often from complex network analysis: short average path length, high level of clustering, and power law and exponential degree distributions (Albert & Barabasi 2002b, pp. 48-49). A short average path indicates a short distance between nodes in a network, while topologically close nodes that are well connected form clusters. In 1998, Watts and Strogatz formalized this cluster concept for complex networks using several large data sets. The real world networks they analyzed were not completely random but instead displayed clustering at the local level. Local clusters linking to other local clusters formed "Small worlds." This analysis was followed by studies performed by Albert and Barabasi (2002b) and Adamic and Huberman (1999), which concluded that when the WWW is studied as a graph it follows power law distribution[2] rather than Poisson or exponential distribution.

Albert and Barabasi (2002b) have described research of the Internet in two realms; the World Wide Web and the Internet. Albert and Barabasi (2002b) label the documents (web pages) of the Internet as the nodes and hyperlinks (URLs) as links. Lawrence and Giles (1998,1999) have estimated the size of this network as having close to one billion nodes based on 1999 data Network research of the WWW has increased as the network experienced rapid growth, and after it was realized that the distribution of the web pages "followed a power law over several orders of magnitude" (Albert & Barabasi

[2]A power-law implies that small occurrences are extremely common, whereas large instances are extremely rare. A function, f(x), is a power law if the dependent variable, x, has an exponent (i.e. x is raised to some power).

2002b, p. 49). Albert, Jeong, and Barabasi, (1999), Lawrence and Giles (1998, 1999), Adamic and Huberman (2000), and Adamic (1999) are a few researchers among the many that have studied the complex network topology of the WWW and Internet. According to Albert and Barabisi, the topology of the Internet is studied at two levels: the router level and the interdomain level (Albert & Barabisi 2002b, p. 49). All nodes are routers and all links have physical connections between them at the router level. The interdomain level consists of hundreds of routers and computers, each represented by a single node (Albert & Barabisi 2002b, p. 52). The interdomain level and the router level have both been studied by Faloutsos et al. (1999), who concluded that in each case, the degree distribution follows power law. The connectivity of the routers was mapped by Govindan & Tangmunarunkit (2000). Yook et al. (as cited in Albert & Barabisi 2002, p. 52) and Pastor-Satorras et al. (as cited in Albert & Barabisi 2002b, p. 52) have confirmed in their studies of the Internet that the network does display clustering and small path length.

The majority of research on complex networks revolves around abstract or theoretical networks and geography is not relevant. But networks do impact geography, and vice versa. At the same time, the Internet is dramatically affecting the city, making research of the Internet's geography relevant and important. The following research will contribute directly to geographic literature and research of the Internet backbone network, the effect of the Internet upon cities, and the study of complex networks.

Modeling Networks with Geographic Information Systems

A geographic information system (GIS) can be used to model a network. There are various GIS structures that can be used as tools in modeling linear features, including coverages, geodatabases, geometric networks, logical networks, optical networks. GIS based modeling programs tie linear features to spatial coordinates, unlike other modeling platforms.

Network modeling is dominated by vector GIS, though it is possible to model most networks using raster-based GIS (Zeiler 1999, Malczweski 1999). Bernhardsen (2002) discusses raster connectivity operations, which is a process that requires discrete cell-by-cell displacements, that originate from a single starting point. The cells must contain values that are significant in how one can move on the surface. This means the raster cells represent a friction surface. It is easier to model path attributes such as direction and flow in a vector GIS. The grid cells used in raster only approximate the exact shape of a line in a network, direction is not explicitly given, and line and node attributes must be stored as a separate layer (Bernhardsen 2002).

GIS based systems enable the user to take advantage of dynamic segmentation. This is an extremely important feature in building a network model. Dynamic segmentation is a two-step process performed on a spatial data set comprised of linear features. First, a route system is created by associating adjacent line segments into one or more groups that have a definite linear sequence. Second, descriptive information is associated with the route system by referencing distances from the starting point of each route. Dynamic segmentation allows tiny areas along a line feature to be referenced without actually breaking that line into smaller pieces. This means that linear distances can then be calculated directly from the routes and associated attributes (Northwest GIS Services 2002). Dynamic segmentation uses a linear referencing system (linked to geographic coordinates) to define a common datum for referencing the linear lines (Zeiler 2002).

O'Sullivan and Unwin (2003) explain that while software packages such as ESRI's Network Analyst are showing great promise in the realm of network analysis, a complete comprehensive tool kit that will address the complexity of line objects and the advanced mathematical concepts needed for analysis is still years away. This is because statistical approaches to lines, as well as graphs, have had only limited success. In agreement,

Malczewski (1999) notes that some researches contend that there are operational limitations on the use of optimization models for spatial decision analysis in a GIS environment. But, Malczweski maintains that although GIS presently optimizes in data gathering and visualization of the results, it can be fully integrated to provide a powerful tool for spatial decision support in multi criteria decision making.

GeoDatabases and Geometric Networks:

Geometric networks are networks that model linear systems such as utility networks and transportation networks (MacDonald & ESRI 2002). They support a rich set of network-tracing and solving functions. Geometric networks consist of edge network features and junction network features (Zeiler 1999). Edge elements are connected to other edge elements via junctions. There are two types of network features, simple and complex. Simple network features correspond to a single network element, while complex features correspond to more than one network element.

Principal benefits to the geometric network model (Zeiler 2002, p. 128) include the following:

- Editing networks is simple. When a user adds network features, one can ensure that they are properly connected to the rest of the network with network connectivity rules.
- Network features can represent complex parts of a network, such as switches. This simplifies the editing process and allows one to create maps of a higher quality with less features in one's network representation.
- A suite of simple and advanced network analysis solvers is built into ArcInfo, ready to use. Network analysis is fast even on very large datasets.
- Networks can be versioned. Multiple users can simultaneously edit the same large network in compliance with their organization's work-flow practices.

Geodatabases, part of ESRI's ArcGIS software is a unique data format that is similar to the coverage data model. It is a storage mechanism for spatial and attribute data that contains specific storage structures for features, collective features, attributes, relationships between attributes and relationships between features.

There are two main concepts to understanding a geodatabase:

- A geodatabase is aphysical store of geographic information inside a database management system.
- A geodatabase has a data model that supports objects with attributes and behavior. Behavior describes how a feature can be edited and displayed. (ESRI 2002).

A geodatabase has the capability to allow multiple users working from it simultaneously: Geometric networks are created using geodatabases. The data and network functions and flows and relationships are used to build a geometric network model through the geodatabase. Given the capabilities and sophistication of the geodatabase and geometric network models, it would be ideal to build a telecommunication infrastructure data model using these tools given flow or line data are available. The geometric network model allows several data types to be incorporated into the model, which is what the current project calls for. Several types of telecommunications data with different characteristics and capabilities are being studied. The model would allow not only for organization and a model of the data, but simulation exercises and practices that would not be possible in non-GIS supported network modeling environments.

Geographic Literature of Networks & Information

Geographic Network Analysis

The Internet is a data/information transport network with the ability to connect places that are geographically separated, moving data from node to node, user to user, service to service, workstation to workstation. Though geographers have a long history of applied network analysis (e.g., Lalanne 1863), relatively little has been done on the geography of the Internet. Geographic research has mainly focused on transportation networks. In 1961 Garrison and Marble published their research findings on the U.S. transportation system in The *Structure of Transportation Networks*. They concluded that

transportation structure is dependent upon the characteristics of the location housing the network. Garrison and Marble (1961) also incorporated the work of fellow geographer Brian Berry (1960) into their analysis by utilizing his measurements of technological and demographic factors. Berry had done substantial research on networks and economic variables, incorporating technological and demographic variables into his research. He synthesized statistical measurements of levels of development to reveal the basic factors underlying variations in the measurements of development. Partnering with Berry, the researchers were able to incorporate national development into their analysis using regression methods. They found that technological development was the major determinant of network structure and that physical characteristics of a location are less significant in explaining network structure than level of development of a location (Taaffe & Gauthier 1973, p. 112). Kansky continued the research of Garrison and Marble (1961) and Berry (1960) in his 1963 paper titled "The Structure of Transportation Networks." Adams (1971) followed the U.S. highway network research with an analysis of the domestic airline network. Adams used matrix methods to study airline growth and connectivity. Nyusten and Dacey (1968) expanded the research agenda to include telephone networks & other types of fixed infrastructure. Taaffe and Gauthier (1973) followed with a text that demonstrated and explained how geographers study transportation systems. In 1977 Haggett, Cliff and Frey published two volumes explaining locational analysis & methods that are still applied in geography today. The major contribution of this research was the development of spatial models for network structure relating to location, density, and change over time. Haggett, Cliff and Frey also explored network nodes, and the hierarchical structures they form within networks.

Geographers have also contributed to network analysis in related fields. Mitchelson and Wheeler (1994) illustrated the importance of information flows through the U.S. in terms of the global economy. Longcore and Rees (1996) have shown the importance of

telecommunication infrastructure in financial districts. Hepworth (1990,1991) has studied the Geography of the Information Economy He concluded that IT convergence has led to the centralization of information activity while communication between locations has enabled the decentralization of knowledge-creation.

Using FedEx geographic delivery data, Mitchelson and Wheeler (1994) explored the relationship between information flow and the U.S. city hierarchy within the context of the global information network. They defined criteria for unstable economic conditions: deregulation, globalization, demassification, and vertical disintegration as important factors in promoting instability. These conditions are dependent upon the information economy, and the exchange of information is critical when instabilities arise The Internet is a tool to transfer information, data, and ideas, and acts as a stabilizing force in the global economy. Mitchelson and Wheeler's analysis can be applied to the Internet to give insight into information flows and spatial structure of information economy, just as the FedEx delivery system is used to establish a domestic hierarchy. Longcore and Rees (1996) built upon the work of Moss (1991,1998), Castells (1989, 1993, 1996, 1997, 1998, 1999), Sassen (1991, 1994, 1995, 1996, 1999, 2000), Dicken (1994, 1998) and others to study information technology and networks at the local level. Longcore and Rees (1996) used the financial district in New York City to assess inter-urban information flows. The found that the decentralization of central city office activity was enabled by electronic communications and concluded a new urban hierarchy was emerging, based on inter-urban information flows (Longcore & Rees 1996). However, they also concluded that only global cities could support a sufficient concentration of telecommunication infrastructure. The demand for telecommunication infrastructure would exist in larger cities, implying that infrastructure capacity is reflective of underlying market conditions and position in the global hierarchy.

More recently, the information and economic flows of e-commerce[3] have been studied. Leinbach and Brunn (2001) address the rapid growth of IT[4] sectors that demand a technically trained and highly skilled workforce. They conclude that the major cost burden of communication infrastructure has fallen on the shoulders of the private sector. Button and Taylor (2001) and Kenney and Curry (2001) illustrate the importance of the Internet and e-commerce. They both conclude that the Internet is an important tool for reducing transfer costs. Goodchild (Leinbach & Brunn, 2001) addresses the location theory implications of the Internet and e-commerce and concludes, "The Internet is more than just another communications device. It is a newly developed space wit the power to give rise to novel forms of human social interaction in almost any area of human endeavor, commercial, or otherwise" (Goodchild 2001, as quoted in Leinbach & Brunn 2001, pp. 63-63). Malecki and Gorman (2001) study the physical structure of the Internet and the importance of geography asserting that the Internet illustrates "both old and new geographies" (Malecki & Gorman 2001, as quoted in Leinbach & Brunn 2001, p. 103). Still others have studied e-commerce in firm, regional, and global contexts: Aoyama (2001), Cobb (2001), Coe and Yeung (2001), von Geenhuizen and Nijkamp (2001) and Langdale (2001).

Cukier (1998) tackles the geography of the Internet on a global scale, concluding that in a postmodern world of consumerism and industry, geography matters; but in a digital economy, information is the main product of value, and connectivity is what really matters. Recent claims have touted the "death of distance" (Cairncross 1997) in business. President Bill Clinton's 1998 address to the United Nations proclaimed the Internet is responsible for "the death of distance," and he asked the United Nations to support the new technology [the Internet] (Clinton 1998). Dodge and Kitchin (2001) have

[3] E-commerce refers to the exchange of goods and services via the Internet.
[4] IT refers to Information Technology

continued to stress that geography remains important despite these claims. They detail a literal, conceptual, and metaphorical mapping of information and communication technologies and cyberspace and conclude that even in cyberspace, geography matters. Warf and Purcell (2001) contend the idea that the relevance of geography and location are pertinent because the Internet compiles and portrays a definite spatial structure that reinforces existing relations of wealth and power. They acknowledge that "though deregulation and digitization have severely attenuated the linkages between money and space" global money does not "presuppose the disappearance of the nation-state, but rather a rearticulation of its functions" (Warf & Purcell 2001, p. 240).

When users are browsing the web, traveling from site to site and location to location, geography is, to the user, of little relevance. However, the geography of the Internet's infrastructure is of great relevance. Wilson (2001) reminds us that seeking territory in cyberspace has both "metaphorical and real geographic elements." Wheeler, Aoyama and Warf (2000) have produced a publication that concentrates on the geographic distribution of telecommunications and discuss how changes and innovations in the economic system are catalyzed by telecommunication networks. Their publication includes descriptions of how telecommunications have brought about the restructuring of cities such as Atlanta, Phoenix, and Sunderland, England. They cover the geography of Internet real estate, telecommuting, and urban planning and attribute changes in the economic system to the heavy influence of telecommunication networks.

Recently there have been two ideas about the effects of telecommunications on cities. The first idea is that information transfer will replace distance, causing the death of cities (Gilder 1995). "Some social theorists argue that new information technologies will inevitably lead to the economic decline of cities as electronic communications make it possible to replace the face-to-face activities that occur in central locations" (Moss 1998). "We are headed for the death of cities" (Gilder 1995). The second idea is a little

more rational, given that cities continue to experience population growth: Telecommunications technologies are not a replacement for personal interactions, but an enhancement.

> It is also possible that telecommunications are not a substitute for face-to-face interactions, but in fact these two forms of information transmission are complements. If they are compliments, then we should expect cities and [selected urban] space to get more important as information technology improves. (Moss 1998)

The implication is that telecommunication infrastructures are likely to reinforce existing trends rather than create divergent trends.

The analysis of transportation and telephone networks has been an important research topic in geography for some time. Nonetheless, there has been an absence of studies on the Internet and Internet infrastructures. Geographic analysis of the Internet has increased in recent years, though mostly describing the growth of Internet hubs and capacity and the geographic distribution of networks and the Internet's users and traffic. Little emphasis has been placed on the complex connectivity of the Internet from a network analysis standpoint, particularly the Internet backbone network.

Geographic Research of Telecommunication Infrastructures

Geographers have begun to analyze the Internet and telecommunication related infrastructures; including colocation facilities, Network Access Points (NAP), Metropolitan Area Exchange's (MAE), Internet Exchange (IX) Points, Marine cablelandings, Point of Presence (POPs), Internet backbones, fiber routes, cellular towers. Recent work has also examined wireless structures (Gorman & McIntee 2003); Web content; and information production and distribution on the Internet has also been explored (Zook 2001, Wilson 2002); and the locational attributes of colocation facilities (McIntee 2001).

The colocation industry emerged as demand for the interconnection of telecommunication networks rose dramatically with the growth and proliferation of the

Internet. These facilities serve as physical interconnection hubs for ISPs, Internet backbones and servers and are known by many nicknames: telehouses, telecom hotels, and Internet hotels. The location characteristics of other types of Interconnection hubs, NAPs, MAEs, IX Points, and marine cable landings have also been examined (McIntee 2001). These interconnection facilities are clustered in cities rich in telecommunication infrastructure, specifically fiber-optic networks, connecting the networks. Evans-Cowley, Malecki and McIntee (2002), Malecki (2002), and Malecki and McIntee (2003) have further explored the colocation industry and the geographic location of these facilities and their effect on urban places and urban structure. Telegeography (2001b) also has contributed to information of the geographic distribution of the colocation industry by compiling datasets of colocation facilities.

Point-of-presence (PoPs) facilities are a type of infrastructure that allow for the connection between local Internet service providers and Internet backbones. Grubesic and O'Kelly (2002) found that the greater San Francisco area led all U.S. metro areas in the number of POPs in 2000, likely attributed to the high concentration of Internet networks housed in the city

Another important telecommunication infrastructure that has been studied by geographers is cell towers. Gorman and McIntee (2003) found a strong and significant relationship between cell towers and the volume of data traffic and the location of colocation facilities within a C/MSA implying that market size and urban growth are key in the location of cell towers. In 2001, 65,000 cell towers existed in the U.S. Of that number, 41,204 were located in C/MSAs (Gorman & McIntee 2003).

Web content, hosting facilities, and location of information production have also been studied in geographic contexts (Malecki 2002, The Economist 2001). Zook (2001) has explored the physical locations of adult video content providers vs. online content providers. He concluded that there is a "stronger connection between Internet content

and information–intensive industries than between the Internet and the industries providing the computer and telecommunications technology necessary for the Internet to operate" (Zook 2000, pp. 411-412). Wilson (2002) has analyzed the geographic location of virtual casino domains. He found that in 2001, the U.S. led the distribution of casino domains, housing roughly 25% of the world's casino sites; but the urban distribution within the U.S. was widely dispersed compared to the Internet industry at large.

Considerable research has been done into the affects of new communication technologies on cities and the urban hierarchy. These studies revealed that communications infrastructure has disproportionately agglomerated in the largest metropolitan regions (Malecki & Gorman 2001, Malecki & McIntee 2000, McIntee 2001, Moss & Townsend 1998, Wheeler & O'Kelley 1999), a pattern that reinforces the predominance of those metropolitan areas within the urban hierarchy. This concentration of infrastructure in the largest cities confirms the theory that telecommunications infrastructure and technology will not bring the decline of cities, but rather complement cities in their attempt to stay viable as centers of commerce. Although past research on telecommunications includes the networks that comprise the Internet, little is presented other than the topology of the nodes and links (Hepworth 1990, Kellerman 1993, Malecki & Gorman 2001). Topics of importance are included in Brunn and Leinbach's (1991) *Collapsing Space and Time: Geographic Aspects of Communication and Information*: geography and communications, information economies, communications, technologies, and regional development, and social dimensions of information and communications. Longcore and Rees (1996) studied city structure change, influenced by the most recent changes in information technology, using Manhattan as a case study (Longcore & Rees 1996). Although their research concluded that the financial district land market might cause the tightly focused financial district to demonstrate geographical flexibility, they

recognize the importance of face-to-face contact, and proximity to sophisticated telecommunications infrastructure (Longcore & Rees 1996).

Geographic Research of the Internet Backbone

Within the past decade various researchers & organizations have looked at the Internet backbones. Most of this research has focused on the Internet backbones in the U.S., though some has discussed the structure of the Internet in Europe. The following section reviews geographic literature of the Internet backbone network.

Telegeography (2000, 2001) reported that the international backbones with the largest bandwidth capacity in 2000 were not surprisingly located between London and New York (26680.5 mbps) and between London and Paris (24340.5 mbps[5]). The backbone link between San Francisco and Tokyo had the largest bandwidth capacity between North America and Asia in 2000 (capacity 7550.0 mbps) (Telegeography 2000, 2001). Telegeography (2001) has also reported that New York serves as the Internet's most global metro area, directly connected to 71 countries in 2001. Five of the top-ten cities cited in that report were intercontinental backbones located in the U.S. This reinforces the theory that Internet traffic is heavily reliant upon the U.S. as a centrally located switching hub in the global telecommunications network.

Between 1997 and 1999, the U.S. experienced large, rapid growth in the Internet backbone network with a 420% increase in data transfer capacity (Moss & Townsend 2000). Moss and Townsend (2000) also reported that they found an increasing concentration of Internet backbones in several mid-sized locations that were centrally located. World Com, Sprint, and Cable & Wireless dominated the Internet backbone network in 1999, controlling about 55% of the domestic market (Telegeography 2000). By 2001, with mergers and acquisitions flooding the market, World Com controlled 37% of the domestic market (O'Kelly & Grubesic 2002). O'Kelly and Grubesic (2002) found

---

[5] Mbps- Million bits per second. A measurement of data transfer rate

that East coast cities in the U.S. experienced a high concentration of Internet bandwidth, a phenomenon that begins to slowly diffuse westward. They attributed the high concentration of links and bandwidth in Washington, D.C. to the combination of its role as a capital city and it's high-tech industry. They also concluded that Chicago was the most- accessible city in the U.S. backbone network based on the fact it had more Internet connections or pathways between it and every other U.S. city.

Wheeler and O'Kelly (1999) studied the accessibility levels of 31 Internet backbones in 1997. They found that Washington, D.C., Chicago, San Francisco, New York, and Dallas led the U.S. in the most accessible cities in the Internet backbone network. Malecki and Gorman used connectivity matrices to determine U.S. city hierarchies based on 1-hop links and 2-hop links in the bandwidth-weighted matrix. They concluded that network analysis of the Internet illustrated old and new geographies; the Internet has changed the meaning of distance, space and the geographical significance of places, following old routes while also establishing new ones (Malecki & Gorman 2001 in Brunn & Leinbach, p. 103). The use of binary connectivity matrices confirmed the "strong spatial bias and hierarchical structure of U.S. cities-one that differs from the conventional population-based hierarchy" (Malecki & Gorman 2001 in Brunn & Leinbach p. 103). Malecki and Gorman found that the major cities in the economy double as the major nodes of the Internet.

History of the Internet and Parallel Developments in Telecommunications

Early Communication Networks

Communication networks date back to antiquity. The story of Phidippides, who ran in 490 B.C. 36.2 km from Marathon to Athens to warn the Athenians of an approaching army is one of the earliest examples of a communication network (Holzman & Pearson 1995, p.1). Early communication systems included the Pony Express (1860-1861), pigeons (as early as 776 B.C.-still used in 1981 by an engineering group in California),

mirrors and flags, fire beacons, watchman and sentors. The first telegraphic device was reportedly around 350 B.C., when a rudimentary device using fire signals to direct flow of water in Italy. There followed a two-thousand year gap in telegraphic devices until the telescope as invented in 1608 by Hans Lippershey (Holzman & Pearson 1995, p. 31). More modern telecommunication closely related to today's communication networks arrived in 1844 with the invention of the telegraph (Hugill 1999, Wilson & Corey 2000).

A Brief History of Telecommunications

The telegraph is described as the earliest ancestor of the Internet; like most communication technologies such as the telephone and the Fax machine, the Internet has been built upon the foundation of the telegraph (Standage 1998). The telegraph was the first in a long series of inventions and technologies designed to exchange information electronically (Lebow 1995) The printing press cannot be overlooked, however, as it allowed the first type of "one-to-many" communication and introduced a mass-produced format that allowed for fairly rapid exchange of information and data. Telecommunications changed little until networked computers allowed "many-to-many" communications (Malecki 2002). The many-to-many communication has been both an aid and hindrance. For example, users are able to e-mail multiple recipients with news or information. At the same time, users are subject to annoying e-mail and advertisements coined as "spam."[6] The media are constantly informing us that we are in the midst of a communications revolution due to rapidly changing information transfer technologies. It may be relevant to acknowledge that the electric telegraph was a far more disruptive technology to its era than the Internet has been to us (Standage 1998). The printing press and telegraph are most credited to having an impact as significant as the Internet: the printing press and the telegraph (Malecki 2002).

---

[6] Spam is the term used to describe unsolicited "junk" e-mail sent to large numbers of people to promote products or services.

Two notable differences in telecommunications and transportation are apparent since the arrival of the telegraph: moving intangible goods, data, and information are not the same as moving tangible goods (Hillis 1998). In addition, telecommunications is to function as a network with simultaneous utilization by many users sending and receiving such intangibles (Rosenberg 1994), making telecommunications a great influence on business. In short, business began to use the new technology to exchange information without involving physical movement of man or animal to do so. Banks and financial corporations were the first types of businesses to take advantage of this new technology (Beniger 1986, Gabel 1996). Telecommunications revolutionized interaction between individuals and institutions as well as created a new platform for networks and networked systems.

Communication technologies have become increasingly bundled in recent years. Many types of communication devices are being developed to serve multiple purposes with maximum convenience. Kellerman (2002, p. 15) describes how it has become possible to use the computer as a telephone, fax and TV, and receive several of these services from a single service provider. He muses that "this fusion may possible mature into a single appliance for information consumption and production, as well as so-called public networks of data and software" (Halal 1993, as cited in Kellerman 2002, p. 15). Phone companies such as Sprint and Verizon are now offering cellular phones that have the capability of wireless Internet access, fax, photos, and more (www.sprint.com, www.verizon.com). Phone companies are also offering direct service lines (DSL) in addition to regular phone service. Cable companies have also expanded their services to compete in the new high-speed Internet access industry. Media conglomerate Time Warner's cable division recently introduced 'Roadrunner' to compete in the Internet service provider market. Roadrunner is a "high-speed, online service providing unique

broadband[7] content, services, and lightning-fast access to the Internet. Road Runner is delivered to your computer over the same upgraded cable systems that currently bring cable television to the home (http://www.roadrunner.com). Today, many industry analysts predict that with the growth of data networks, voice traffic will increasingly travel over Internet protocol (IP) technology. With increasing data traffic, the demand for Internet fiber is on the rise (National Research Council 2001).

Telecommunications and the City

Location Decision and Telecommunications

Telecommunications technology has been booming since the early 1970s. Prominent advancement and change in telecommunications technology have influenced location decisions in business. Different types of firms are scrambling to locate in areas rich in technology infrastructure. Traditional location theory typically includes: local input and output, transferable inputs and outputs, climate, labor supply, taxes, and local economy. Many firms continue to use traditional location factors but are beginning to incorporate new factors into location decision, especially high-bandwidth Internet connection. These firms include those involved in banking, research, marketing, telecommunications, and many more. High-bandwidth connectivity is an increasing attractive asset to firms' location decision but at the same time, the firm's location is pertinent. The assumption that a firm can ignore geographic location because of technology is false. Geographic location is still an important location factor for many firms. Firms dependent upon technology search out locations that have strong technological infrastructure. This infrastructure might include: switches, POPs (point of presence), NAPs, Gateways, etc.

---

[7] Broadband describes a transmission facility having a bandwidth sufficient to carry multiple voice, video or data channels simultaneously.

The presence of modern telecommunication technology reduces transfer and information acquisition costs. Transfer of data is dramatically reduced in terms of time. Transfer costs of physical commodities and production inputs for assembly have traditionally been a main factor of location decision, and still are. However, the actual transfer of some goods has changed given the switch in emphasis to information services and intangibles. The transportation of data, electronic mail, music, movies, and news are not physically transferred. These goods can be transferred electronically. Other goods (such as food, clothing, bicycles, and tangible goods) must be physically transported. Many firms use both types of transfer. For example, a company that sells women's apparel. This company might be involved in e-commerce, but the final clothing item must be shipped to the customer. So geographic location continues to be an important factor in this firm's location. The firm must be "connected" technologically, while at the same time able ship goods to customers at minimum costs. Telecommunication networks are "friction reducing" technologies, that enable transfers between remote locations for costs that are substantially lower than the physical transfer of information between them" (Salomon 1988) via human interaction or hard copy, etc. Whittaker Associates has identified the most important site-selection factors to the Business Services industry. Of the 51 factors studied, the top-ten include the following:

- Telecommunication services
- Secondary education quality
- Effective cost of skilled labor
- Effective cost of unskilled labor
- Availability of executive, administrative, managerial workers
- Administrative support
- Geographic proximity to markets
- Access to business & tech. Services
- Business taxes
- Energy dependability

Telecommunications technology will increase the potential of cities, and has in fact, since the 1980s, revitalized the central business districts of the leading cities and

international business centers of the world-New York, Los Angeles, London, Tokyo, Paris, Frankfurt, Sao Paulo, Hong Kong, and Sydney. These cities have reached their highest density of firms ever, providing further evidence that cities are not on the decline (Sassen 2000). With the knowledge of the benefits sophisticated telecommunication technology can bring to urban centers, many cities are welcoming infrastructure. Sophisticated technology can benefit the city by attracting those firms seeking the infrastructure, such as financial, media, and web-based firms, which in turn provides jobs, education, and services for those who have access to the infrastructure. Those who do not have access to the technology are at a disadvantage. Telecommunication technology is changing the economies of networks, emphasizing service industry rather than manufacturing.

Many firms have included technological infrastructure as important factors in their location decision factors. The highest level of advanced telecommunications infrastructure is found in the country's largest cities; implying that cities that are better connected with sophisticated telecommunications infrastructure are a more attractive location. These high urban centers are able to cash in the technological amenities they offer. Many public services such as libraries, tax and finance administrations, and criminal justice systems are information intensive, dependent upon computers, telephones, and sophisticated information retrieval and imaging systems. "A city's future as an information center depends on information-producing activities that occur through both face-to-face and electronic communications" (Moss 1998).

Infrastructure in Cities

The intra-urban patterns of telecommunication infrastructure are greatly dependent upon each other. For example, the physical structures of the fiber-optic networks on the ground are greatly dependent upon interconnection facilities. At the same time, colocation facility location is just as dependent upon the concentration of fiber networks.

The same cities that lead bandwidth rank also lead colocation rank (McIntee 2001). Cities that lead the U.S. in terms of bandwidth concentration consecutively lead the U.S. in terms of colocation facility concentration. The relationship between different types of Internet infrastructure is a simile to the figure of speech "which came first, the chicken or the egg?" The relationship between colocation facilities and Internet bandwidth is no exception. It is difficult to determine which is more dependent upon the other, as colocation facilities are interconnection points for Internet backbones, while at the same time, colocation facilities locate in close proximity to termination points of Internet backbones.

The Internet has sparked a concern for new legislation and policy to help protect those who use the Internet as well as those who are affected by the Internet and its infrastructure. There are many concerns with the effects of the Internet, and its infrastructure location within metropolitan areas. The negative effects can be blanketed under one term. the digital divide (Sassen 2000, Wilson & Corey 2000, Wheeler, Aoyama & Warf 2000). The digital divide is a simple definition for the gap between the rich and the poor widening as those with wealth and affluence have access to the increasingly valuable advantages of sophisticated technology, such as the Internet, while the poor are further disadvantaged because they do not have equal access to this valuable tool.

Those firms that benefit greatly from Internet infrastructure especially those involved in financial services, media, consulting, are using infrastructure as an increasingly important factor in location decision (Finnie 1998, Kotval 1999, Longcore & Rees 1996). Communities have been using their telecommunications infrastructure as a strategy to attract new business and to increase their overall economic competitiveness. In a European survey of 500 companies, telecommunications was cited as the second-most important factor (Graham & Marvin as cited in Kotval 1999). Cities such as New

York, Boston, and Amsterdam have earned a competitive advantage by establishing "teleports" (satellite linkages connecting to local telecommunication networks), and they effectively result in a globally networked city. For those who do not have access to the technology, the ever-growing sophisticated infrastructure could create problems. Universal access to information and communication technologies is critical in closing the gap between the economically disadvantaged social groups and the advantaged groups. Graham (1999) is already describing the unfavored zones within cities as "network ghettos," places of low telecommunications access and concentrated social disadvantages. "Uneven global interconnection via advanced telecommunications becomes subtly combined with local disconnection in the production of urban space" (Graham 1999). Those who could benefit most from the infrastructure as a tool to enhance quality of life, job searches, education, and communication, may be those who are excluded from the sophisticated technology.

Poor and less-advantaged cities that are reluctant to welcome Internet infrastructure and telecommunication competition may be disadvantaging themselves. Finnie (1998) studied 25 major cities, and determined that global cities that remain competitive in attracting business firms lack strict regulation within the telecommunication sector. According to Finnie, telecommunications services are becoming increasingly central to business success or failure, as competition increases and sophisticated technology becomes more widely available. As a result, the gap between the haves and the have-nots could well be narrowing. Because of all the benefits cities receive when Internet infrastructure is implemented, it is not feasible to ban growth of these sophisticated networks within urban areas.

Kotkin (2000) has claimed that the digital era we are currently experiencing is a period of advancement not seen since the industrial revolution. It is hard to argue this statement. The dawn of the Net has impacted the economic and social geography of

America largely, and some believe it is redefining the American city hierarchy (Kellerman 2002, Kotkin 2000, Townsend 2001). Numerous titles have been devoted to the digital revolution, the Internet, and an increasing interconnectedness of our world. *Six Degrees* (Watts 2003) and *Linked* (Albert & Barabasi 2002a) discuss the overlap and interconnection of networks in modern day life. Some texts, such as *Information Tectonics* (Wilson & Corey 2000), *The New Geography* (Kotkin 2000), and *Worlds of E-Commerce* (Leinbach & Brunn 2001) explore the economic, geographic, and social implications of the digital revolution from the perspective of various disciplines.

There is growing literature on the Internet, it's history and composition, and the effect of proliferating telecommunication networks on cities, complex regional networks. This research helps to provide a perspective for examining the U.S. Internet backbone network. This dissertation will contribute to that literature by analyzing the connective properties of the U.S. Internet backbone network.

Data and Methodology

Internet backbone data have been obtained from George Mason University School of Public Policy. The dataset was created by researchers at George Mason University in 2003. The Internet backbone dataset provides a measure of the amount of data capacity and connections a consolidated metropolitan statistical area (C/MSA) has to move information to another C/MSA. The data was calculated from the total long haul fiber capacity, or bandwidth, connecting a C/MSA to other C/MSAs. Bandwidth is the term used to describe transmission speed, which is measured in bits per second. According to Malecki and Gorman (2001), "bandwidth is what makes communications— specifically Internet Protocol (IP)—different from transport networks. The limiting factor of IP networks is not distance, but the capacity of the bandwidth available on the network from one location to another" (p. 90). The normal speed of a voice call is 64 kbps. Transmission speeds above 64 kbps is generally categorized as broadband (Huston

1999, pp. 160-171. The Internet backbone providers are companies that own the framework of the Internet. This network connects CMSA nodes and transports data across long geographic distances. Multiplexing is the process of sending multiple signal streams of information on a backbone at the same time in the form of a single signal. By multiplexing, higher bandwidths can be achieved. In 2001, 48 private providers operated the Internet backbone networks. These firms range from large telecommunication carriers such as AT&T, MCI WorldCom, Sprint, IBM, and Cable & Wireless, to smaller, lesser-known firms. The backbone providers are called autonomous systems (AS), which means they operate independently from other systems, setting their own policies and network structure (Malecki & Gorman 2001, pp. 92-93). These independent networks interconnect to form a larger network, thus creating the Internet backbone. In 1995 Huitema suggested that the Internet backbone network is the best indicator of the geography of the Internet. The amount of bandwidth between C/MSAs is not equal. Chapter 3 analyzes the network as though the bandwidth connections are equal, while Chapter 4 address the inequalities by adding weights to the analysis.

A geographic information system (GIS) was created with the Internet backbone network data, using ArcGIS from Environmental Systems Research Institute (ESRI). Several types of telecommunications data with different characteristics and capabilities were also incorporated into the GIS. The GIS were statistically analyzed to determine spatial relationships between the Internet backbone network and to understand the distribution of Internet bandwidth. The telecommunication infrastructure were geo-coded, giving each datum spatial attributes so it can be analyzed in conjunction with other types of data in a common information system. Telecommunication infrastructure data obtained for this research project include the following:

- Telephone Switches (Digital, Wireless)
- Cellular Towers (Towers, Antennas)
- Fiber Lit Buildings

- Colocation Facilities
- Network Access Points (NAPs)/Metropolitan/Area Exchanges (MAE)
- Marine Cablelandings
- Fiber Points of Presence (Pops, termination points)

The following is a description of the data used in the GIS and statistical analysis performed in Chapter 6:

Telephone switches. This database includes the location, capability, and ownership of telephone switches. Description includes details such as which switches are wireless, digital, or integrated services digital network (ISDN). Customers might purchase a single type of telephone switch they are interested in, such as wireless switches. The data is geocoded and can be used in a geographic information system.

Cellular towers. A complete database of cellular towers in the United States including the tower owner and capabilities of the structure. Data is also available on the auction results of metropolitan areas. Auctions provide an economic valuation of regions by private industry for the implementation of a technology. The auction values of a region provide a new insight into how emerging technologies are affecting the urban hierarchy of regions.

Fiber lit building. This is a very extensive database that includes numerous fiber carriers. The description includes carriers, street addresses of termination point, type of fiber, capacity and status of fiber, common language location (CLLI) codes, and geocoded. This data is geo-coded and address matched, determining the building and spatial attributes of the fiber location. This database includes dark fiber, lit fiber, fiber currently in existence, fiber "in the pipe," as well as network expansion details of future fiber locations. Fiber loops are also included in this database as are carrier "lit" buildings and metro fiber routes. (Source Geo-tel 2002).

Internet interconnection facilities/colocation facilities. These are provision network providers with floor space for their network equipment within a secure building. The building is typically equipped with appropriate heating, ventilation and air-

conditioning (HVAC), enhanced fire suppression, electrical connections and diesel-powered generators to guard against commercial power failures. Private companies often choose to interconnect in these facilities, as well, to avoid costly local loop charges and to have the ability to cross connect to their carrier of choice. Companies also collocate to utilize multiple carriers so if their primary carriers' network fails, they can reroute network traffic to the back up carrier already in place.

Network access points (NAP), metropolitan area exchanges (MAE), Internet exchange (ix) facilities, carrier hotels (Geo-tel 2002). These are Internet interconnection *facilities* on a grand scale. Some of these facilities are considered "public" interconnection facilities, while others (MAE) are privately held. Many of these mega facilities were original interconnection points for the early Internet.

Marine cablelandings. This data shows the exact location and termination of the marine cableheads and lists the carriers located in the cable. A description of each carrier's fiber in the marine pipe is also available. The spatial attributes for this data are also included in the database (Source Geo-tel 2002).

Point-of-presence data (POP). This data includes the carrier of the POP and the spatial attributes of the POP. This data set consists of carrier fiber points of presence (POPs) that signify termination of fiber lines that provide connectivity to a location, typically office buildings (Geo-tel 2002).

In addition to telecommunication infrastructure data, a database with descriptive statistics of the metropolitan areas was compiled. The database included population, bank deposits, income, and the local economy's dependence on specific sectors, such as finance, insurance, and real estate (FIRE), as well as other factors that could be used as interactive variables in the model. The descriptive data will be discussed more in-depth in Chapter 6.

The data have been modeled in a geographic information system (GIS), using ESRI's ArcGIS software to study the distribution of telecommunication infrastructure in conjunction with like-kind as well as to study the hierarchy of nodes. Chapters 4 and 5 of this dissertation use the long-haul fiber optic bandwidth data for the analysis performed within. Chapter 4 uses the unweighted long-haul fiber optic bandwidth data for the unweighted analysis. Chapter 5 adds weights to the analysis. Chapters 4 and 5 include maps that display the long-haul fiber optic bandwidth data with other types of infrastructure. However, the analysis for these chapters was performed solely using the fiber network data. The analysis in Chapter 6 analyzes the Internet backbone data in conjunction with various types of telecommunication infrastructure data as well. The procedures performed in Chapter 6 also incorporate the descriptive data for C/MSAs into the analysis.

CHAPTER 3
NETWORK ANALYSIS

Analysis

The matrix-based frameworks for analyzing the overall impact of nodal or linkage distribution on the overall connective properties of a network has not been applied to the U.S. Internet backbone network. This research intends to identify the most critical links and nodes in the domestic long-haul fiber network in the U.S. The research methods contained within this chapter can then be applied to other types of telecommunication networks to answer like-kind research questions.

Examining telecommunication networks as a graph has proven to be highly useful in answering the proposed research questions and testing the hypotheses. This chapter introduces and explains graph theory and matrix multiplication. The methods reviewed here will provide a framework for network analysis and will be used to examine the Internet backbone network in Chapters 4 and 5. The purpose of this chapter is to provide a through explanation of graph theory and matrix multiplication that will be used in the following chapters of this dissertation.

Introduction to Graph Theory

A fundamental question in network analysis is the degree to which the nodes are interconnected. The connectivity of a network is defined by the overall degree of connection between all vertices The degree of connection between all vertices is probably the most important structural property of the network (Taffe & Gauthier 1973, p. 101). This section of methodologies will be based largely on Taaffe & Gauthier's 1973 publication, *Geography of Transportation*. Any network can be represented as a graph

(Haggett, Cliff & Frey 1977, Garrison 1960, Kansky 1963).As spatial structures, networks are extremely complex in nature. This makes networks difficult both to describe and analyze. By simplifying networks, we are able to study their characteristics. When applying graph theory to a network in order to analyze it, it is necessary to model the network in the form of a graph. As the network is simplified in analysis preparation, some of the information about the network will be discarded purposely. Only those pieces of information that are most relevant to analysis when using graph theory are taken into account. Noting this, not all networks should be described in terms of graph or matrix theory. Topological analysis takes into account only interconnection, excluding properties such as shape, direction, and size. When the network is studied as a graph, only the topological properties of the network are considered. The large range of characteristics that might be identifiable with various networks are not analyzed in graph theory.

Graph theory breaks the network into points and lines, in an abstract manner. Although it does not model the real world directly, it provides measurement for some structural properties of "a real-world system if that system is idealized as a set of points connected by a set of lines" (Taaffe & Gauthier 1973, p. 101). In the simplest form, networks can be represented by a series of vertices (representing nodes) and a series of edges (representing links), with a relationship of incidence that associates each edge with two vertices. We know only the presence or absence of connections between nodes are given for each pair of nodes and represented in graph form. There are two ways to measure the described network: (1) a single number (2) a vector of number. A single number describes the aggregate geometrical pattern of the network, while the vector of numbers measures the relationship of the individual components of the network to the entire network (p. 101).

There is minimal information given about this network (see Figure 3-1, Hypothetical Network A), so only primitive measures of connectivity can be assumed. The node-link relationships are the only information given to derive conclusions about connectivity.

Figure 3-1. Network A

The first measurements that can be taken are the number of links and nodes. In Figure 3-1, There are 8 nodes(v) and 7 links(e) in the network. Moreover, the network is minimally connected as there is only one link between any two pairs of nodes. Note that there are no redundant links within the network, meaning that no node has more than one direct connection to any other node. Redundances occur when more than one link connects the same two places. With any minimally connected network the number of links is always one less than the number of nodes: $e_{min} = (v\text{-}1) = (8\text{-}1) = 7$. Note that removing any link in this network will disconnect the network into two parts.

Because network connectivity is most meaningful when a network is either compared to another network or used in measuring growth, another hypothetical network (B) is shown as an example for comparison (Figure 3-2).

Network B is more complex than Network A. Network B has 8 nodes and 11 links. This network is more than minimally connected. Most of the nodes in this network are connected to more than one node. When this type of structure exists, the removal of one link will not necessarily disconnect the entire network

Figure 3-2. Network B

In order to compare these two networks (A & B), connectivity measures must be employed. Graph theory provides various simple measures. The most often employed measures include the gamma and alpha indices.

The Gamma Index

The Gamma index is the ratio of the number of edges in a network to the maximum number possible in that network: $\gamma = \frac{actualedge\ s}{\max\ edges} = \frac{e}{e\ \max}$

The number of links in examples A & B can be obtained from counting. There are 7 links in example A and 11 links in example B. The number of possible links (e_{max}) can be computed from the number of nodes in the system. If the network is represented as a planar graph (one where intersections occur only at nodes), the addition of each node to the system increases the maximum number of links by 3. This holds true for any planar[1] network of more than two nodes. Because the graph is planar, the intersection of new links is not possible without the addition of a new vertex. To express e_{max} use $3(v\text{-}2)$. The gamma index then becomes $\gamma = \frac{e}{e\ \max} = \frac{e}{3(v-2)}$.

When using the gamma index to determine maximal connectivity, the relationship between the number of nodes (v) in a networks and the maximum number of links (e), we would use $e = 3(v\text{-}2)$. The gamma index is expressed in terms of a graph-theoretic range that varies from a set of node that have no interconnections, while on the other end of

[1]Planar networks form vertices whenever two edges cross, where non-planar networks can have edges cross and not form vertices.

the spectrum we have a set of nodes in which every node has a link that connects it to every other node in the network. The connectivity "is evaluated in terms of the degree to which the network deviates from an unconnected graph and approximates a maximally connected one" (Taaffe & Gauthier 1973). The gamma index falls between a range of 0 and 1. Using network A as an example, the gamma index would be $\gamma = \frac{e}{e\max} = \frac{7}{3(8-2)} = \frac{7}{18} = 39$. In network B, the gamma index would be $\gamma = \frac{e}{e\max} = \frac{11}{3(8-2)} = \frac{11}{18} = 61$. In terms of maximal connectivity, the first network is 39% connected while the second network is 61% connected.

Alpha Index

When discussing minimal networks, the possibility of linkage removal was discussed. This would sever the connectivity of the network into two separate pieces. Linkages can also be added to a network, thus increasing the connectivity beyond the minimal structure, adding redundancy and/or alternate paths. Additional linkages create circuits. Circuits can be defined as a definite path where the original node of the linkage sequence coincides with the terminal node. If a circuit is present, then it establishes additional or alternate paths in the network. The number of linkages that are added to the minimal network defines the number of alternative paths. The max number of independent circuits in a network is also a function of the number of nodes in the network and the number of linkages necessary for minimal connection between nodes.

The alpha index is a ratio measure of actual circuits, given by (l-n+1), to the maximum number possible in a given network. In a connected network where links are *e* and nodes are *v*, the number of links is equal to one less than the number of nodes ($e=v-1$), only when the network is connected minimally. When there is a circuit in the network, the number of links is greater than the (v-1); $e > v-1$. By subtracting the number of links that are needed for a minimally connected network (v-1) from the actual number of nodes (v), we can obtain the number of circuits in the network. According to Taaffe

and Gauthier (1973), this can be expressed by e – (v-1) = e - v + 1. The resultant is a measure of the number of independent circuits in the network. The maximum number of independent circuits is also a function of the number of nodes in the network and the number of linkages necessary to maintain minimal connectivity between nodes. For a planar network, the maximum number of links is 3(v-2), thus the maximum number of circuits would be: 3(v-2) - (v-1) = 2v – 5. The alpha index is a ratio measure of the number of actual circuits (e-v+1), to the maximum number of possible in a given network (2v-5):

$$\alpha = \frac{\textit{actual circuits}}{\text{max } \textit{circuits}} = \frac{e - v + 1}{2v - 5}$$

The range of the index is from a value of 0 for a minimally connected network, to a value of 1 for a maximally connected network. For the sake of convenience, the numerical value may be expressed as a percentage of circuitry in a network. The alpha values for network A is

$$\alpha = \frac{\textit{actual circuits}}{\text{max } \textit{circuits}} = \frac{e - v + 1}{2v - 5} = \frac{0}{11} = 0$$

The alpha values for network B is

$$\alpha = \frac{\textit{actual circuits}}{\text{max } \textit{circuits}} = \frac{e - v + 1}{2v - 5} = \frac{4}{11} = .36$$

The first network exhibits no circuitry. In the second network, the maximum possible number of circuits is 11, but there are only 4 circuits. The second network's circuitry is 36% of the maximum.

It was mentioned previously that graph-theoretic indices of connectivity are also useful for measuring network growth or change through time. As an example, let us consider that our example networks, A and B are an idealized sequence for transport development. For the purpose of explanation, two more networks will be added. C and D (Figures 3-3 & 3-4). Network C will represent the network prior to A (Figures 3-3). Network D will represent the final network (Figure 3-4). Network C will have less connections than either A or B, as D will have more connections than either A, B, or C (Figures 3-1, 3-2, 3-3 & 3-4).

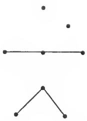


Figure 3-3. Network C

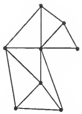


Figure 3-4. Network D

In the beginning stage, which is illustrated by Network C, there are a few links leading to interior centers. In the next state, illustrated by Network A, growth is evident. The network has expanded and includes all of the region's nodes. The growth process continues in Network B (Figure 3-2). Network D is an example of a mature network (Figure 3-4) . The number of nodes has remained constant during the network's growth, but the connectivity of the network has changed. By using the alpha and gamma indices we can determine to what degree the network connectivity has changed, as well as identify the change in the network's spatial structure.

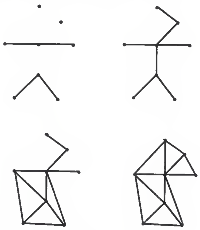


Figure 3-5. Stages of network development

If we arrange the indices in a table, we can see that as the network grows, the connectivity index increases. As the network becomes more structurally complex, both the gamma and alpha indices increase.

Table 3-1. Structural indices for sequence of network development

	γ	α
Stage 1, Network C	.22	-
Stage 2, Network A	.39	.0
Stage 3, Network B	.61	.36
Stage 4, Network D	.78	.63

Three basic network configurations are used to relate the gamma and alpha indices to more specific network characteristics: spinal, grid, and delta. The spinal network shares the characteristics of a minimally connected network, every node is connected to at least one other node and traffic can flow between the nodes but by only a single path. The number of links necessary for a minimally connected network is always one less than the number of nodes in the network (v-1). We can conclude that the gamma index $\gamma = \frac{e}{3(v-2)}$, for a minimally connected network will be $\gamma = \frac{v-1}{3(v-2)}$. The alpha index $\alpha = \frac{e-v+1}{2v-5} = 0$, because there are no circuits in a minimal network. The spinal network is illustrated as $\frac{(v-1)-v+1}{2v-5} = \frac{0}{2v-5}$.

The delta network composition is a stark contrast to the spinal network. The delta network is comprised of a high density of linkages in relation to the number of nodes The delta network composition is one of numerous paths, sequences or links achieving maximal connectivity. The shape pattern most dominant in the delta network is the triangle, for each set of 3 nodes. When a node is added to a network, of more than three nodes, two new links are required. The relationship will always remain 2v-3. Since this relationship will remain constant, the gamma index will be $\gamma = \frac{e}{3(v-2)} = \frac{2v-3}{3(v-2)}$. The alpha index $\frac{e-v+1}{2v-5}$ will always be $\alpha = \frac{(2v-3)-v+1}{2v-5} = \frac{v-2}{2v-5}$.

The third type of network configuration is the grid. The grid network represents the transition network that is sandwiched between the spinal network and delta network. It is a medium between the minimal and maximal network.

To categorize a transport network as spinal, delta, or grid, cutoff values must be established. By determining a scale of alpha and gamma indices for each of the network types, the ranges can be established for each category.

The largest and smallest gamma values are used to identify a spinal network. Taaffe and Gauthier express the gamma index for a spinal network configuration as $\gamma = \frac{v-1}{3(v-2)}$. Alternate expressions given include: $\gamma = \left(\frac{1}{3}\right)\left(\frac{v-1}{v-2}\right)$ and $\gamma = \frac{1}{3}\left[\left(\frac{v}{v-2}\right)-\left(\frac{1}{v-2}\right)\right]$.

Taaffe and Gauthier suggest $\frac{v}{v-2}$ for networks containing a large number of nodes will approach 1 and that $\frac{1}{v-2}$ will approach zero. In sum, this means that the expression will approach 1/3 of (1-0). At the lower end, the value of the gamma index will be 1/2. This means that spinal networks can be categorized between a range of values from 1/3 and 1/2, or .333-.5.

For the delta network configuration, housing maximal connectivity, the gamma index is $\frac{2v-3}{3(v-2)}$ or $\left(\frac{1}{3}\right)\left(\frac{2v-3}{v-2}\right)$ Delta networks will have a range of values between 2/3 and 1.0. Table 3.2 shows the range of values for the three classical networks presented.

Table 3-2. Range of values for the delta index for three classical network patterns

Spinal	$1/3 \leq \gamma \leq 1/2$	*where*	$v \geq 4$
Grid	$1/2 < \gamma < 2/3$		$v \geq 4$
Delta	$2/3 \leq \gamma \leq 1.0$		$v \geq 3$

Considering the network examples A-D, we can see how the network has changed from a spinal network to a delta network. In the first stage of the network, the gamma value is .22. This is expected, as the nodes are isolated, demonstrating minimal connectivity. In the second stage the gamma value is higher, .39. The network is more connected as the gamma value indicates. The third stage finds the connectivity value

even higher, at .61. The fourth and final stage of the network has a gamma value of .78. By this stage, the network displays clearly the triangle configuration that is characteristic of the delta network.

The alpha index can also be used to define network configuration. As discussed previously, the absence of circuits means the alpha index value will be zero. Thus, the alpha value for a spinal network will be zero. The alpha value range for grid or delta networks depends on how many circuits exist. By defining limits of the alpha range it will be possible to determine the network configuration. As defined by Taaffe and Gauthier, the delta configuration for the alpha index is $\alpha = \frac{(2v-3)-v+1}{2v-5} = \frac{v-2}{2v-5}$. The alpha index ranges for the three classical network patterns are presented in Table 3-3.

Table 3-3. Range of Values for the alpha index for three classical network patterns

Spinal	$\alpha = 0$	*where*	$v = e + 1$
Grid	$0 < \alpha < .50$		$v \geq 3$
Delta	$.50 \leq \alpha \leq 1$		$v \geq 3$

The two indices, gamma and alpha, complement each other in network measurement. As a network increases in spatial complexity, the change in indices will be similar. Looking again at the example network, the second stage of the network shows corresponds with the spinal network. This is consistent with the alpha range for spinal networks, which is 0 for the network's second stage. The alpha value for the third stage is .36, categorizing it as a grid network. The gamma value for the third stage is .61, also denoting a grid network. The fourth stage of the network both values fall under the delta configuration. The alpha value is .63 and the gamma value is .78 for the fourth stage. The gamma and alpha indices measure network connectivity and circuitry to describe the network.

Five measures of graph-theory were introduced in the preceding text: number of nodes, number of links, alpha index, gamma index, number of circuits. Diameter is a sixth measure, which has not been defined. Diameter is a measure of the span of

transportation networks, defined as the minimum number of links that are required to connect the farthest two nodes of a network. The diameter describes the minimum number of links required to connect the most distant nodes in a network. Hence, a diameter of five would indicate that there are at least five links separating any two nodes in a network. For example, in network B (Figure 3-2) the diameter is 4. There are only four links separating any two nodes in Figure 3-2. While these indices are useful descriptive tools, it is necessary to remember that graph theory does not include many complexities in practice. In short, graph theory simplifies a network and analyzes it's topology.

Connectivity Matrix

By measuring the accessibility of a node we can determine the hierarchy or the system of competition that may exist between the nodes in a given network. The addition of linkages or the destruction or removal of nodes or linkages is to affect the entire network. Changes of this type also reflect changes in the accessibility or hierarchy of nodes. Graph theory is used to measure these changes as well as to determine if a hierarchy or system of nodes exists as defined in terms of connectivity. Just as any network can be represented as a graph, a matrix can also represent any network. By representing a network as a matrix, numerous questions concerning the network's accessibility can be answered.

Traditionally the origin nodes of a network are represented in the horizontal rows of a matrix and the destination nodes are represented in the vertical columns. The number of rows and columns in a matrix must be identical. Relationships between the nodes are represented by corresponding cells. The points (nodes) of the graph are labeled and the labels are used to identify both rows and columns of the matrix. When two points in the network are connected, this link is represented by placing a non-zero number (typically a

value of 1 in the case of a binary connectivity matrix at the intersection of the relevant row and column. If there is no connection, a zero is placed at the intersection. Figure

Table 3-4 represents a network as a graph and it's matrix format. The number of nodes in the network illustrated is represented by both the number of columns, and the number of rows in the matrix. The presence of non-zero number represents a link between corresponding nodes and the absence of a direct link is represented by a zero. Also, the connection of a node to itself has no value, so a zero is recorded in the corresponding cells. For example, the cell at the intersection of row 2 and column 2 contains a zero. This matrix only gives information on the presence or absence of a direct connection between nodes.

Table 3-4 Example of a binary connectivity matrix

CMSA	Albuquerqu	Atlanta	Birmingham	Boise	Buffalo	Charlotte	Chicago	Cincinnati	Cleveland	Houston	Kansas City	Las Vegas	Los Angeles	New Orleans	New York	Oklahoma	San Diego	San Francisco	Seattle	Tampa	Washington
Albuquerque	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Atlanta	0	0	1	0	0	1	1	0	1	1	1	0	1	1	1	0	0	1	0	1	1
Birmingham	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Boise	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	1	0	0
Buffalo	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0
Charlotte	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
Chicago	0	1	0	1	0	0	0	1	1	1	1	1	1	0	1	0	0	1	1	0	1
Cincinnati	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	1	0	0	1
Cleveland	0	1	0	0	1	0	1	1	0	0	1	0	0	0	1	0	0	0	0	0	1
Houston	1	1	0	0	0	0	1	0	0	0	0	0	1	1	1	1	0	1	0	1	1
Kansas City	0	1	0	0	0	0	1	0	1	0	0	0	1	0	1	1	0	1	1	0	1
Las Vegas	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	1	1	0	0	0
Los Angeles	0	1	0	0	0	0	1	0	0	1	1	1	0	0	1	0	1	1	1	0	1
New Orleans	0	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0
New York	0	1	0	0	1	0	1	0	1	1	1	0	1	0	0	0	1	1	1	0	1
Oklahoma City	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
San Diego	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	1	0	0	0
San Francisco	0	1	0	0	0	1	1	1	0	1	1	1	1	0	1	0	1	0	1	0	1
Seattle	0	0	0	1	0	0	1	0	0	0	1	0	1	0	1	0	0	1	0	0	1
Tampa	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1
Washington	0	1	0	0	0	1	1	1	1	1	1	0	1	0	1	0	0	1	1	1	0

Nodal accessibility can be derived directly from the binary connectivity matrix. This would be the simplest form of measurement of accessibility of a node. By summing each row of a matrix, the total value is the number of direct links between the given node and other nodes. The higher the row total value, the more accessible the given node is to other nodes in the network. Taaffe and Gauthier (1973) use an air transportation network to demonstrate the accessibility of a node (p. 119). In their network, both New York and Chicago have direct flights to each of the other cities in the network. This means that the accessibility of New York and Chicago is greater than other cities in the network. The long-haul fiber data acquired for this project will be used to demonstrate a similar example. A random sample of twenty-one metropolitan areas is used.

Table 3-6. Network represented in matrix form

Nodes	V_1	V_2	V_3	V_4	V_5	V_6
V_1	0	1	0	0	0	0
V_2	1	0	0	0	1	0
V_3	0	0	0	1	0	0
V_4	0	0	1	0	1	0
V_5	0	1	0	1	0	1
V_6	0	0	0	0	1	0

The matrix gives information on the presence of links between each given city, for example there is a direct link between Atlanta and Birmingham. There is a "1" at the intersection of Atlanta and Birmingham to signify a link exists between the two cities. The matrix also denotes the absence of direct connections as there is no direct link between Boise and Albuquerque. A value of "0" appears at the intersection of Boise and Albuquerque (Table 3-4). The matrix also tells us that New York is directly connected to eleven other cities in the sample network.

Table 3-7 illustrates a hierarchy of the twenty-one cities, as shown in Table 3-4 in matrix format, of long haul fiber direct links. Atlanta, Chicago, San Francisco, and Washington, DC, are the most directly accessible nodes in the network measured. If we

added the number of direct links for the corresponding row of each of these cites, the total is twelve. Based only on direct linkages, these cities rank at the top of the hierarchy of nodes.

Table 3-7. Hierarchy of network cities

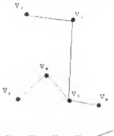
City	Rank	Number of links
Atlanta	1	12
Chicago	1	12
San Francisco	1	12
Washington	1	12
New York	5	11
Houston	6	10
Los Angeles	6	10
Kansas City	8	9
Cleveland	9	7
Seattle	9	7
Las Vegas	11	5
Cincinnati	12	4
New Orleans	12	4
San Diego	12	4
Tampa	12	4
Boise	16	3
Charlotte	16	3
Birmingham	18	2
Buffalo	18	2
Oklahoma City	18	2
Albuquerque	21	1

Much more can be derived from the matrix; however, it is important to note that this measure does have limitations as it only reveals the presence or absence of links and not their capacity to support flows Though a node may have a high connectivity level based on direct connections, it may be lower in the network hierarchy when indirect connections are included in the measurement of accessibility.

Indirect connectivity measures can be counted by using matrix multiplication. This is an element-by-element method that involves multiplying each row of a matrix by the column of another matrix. The sum of the products of the element-by-element multiplication is recorded in the corresponding cell of the new matrix. The first matrix is multiplied by itself; the product of this method is then multiplied by the original matrix to

get the next product. The next product is then multiplied by the original matrix, and so on and so forth.

Figure 3-6 illustrates the matrix multiplication of the matrix introduced in Table 3-4. The matrix is first multiplied by itself (C·C) to produce C^2. For each cell of C^2 the value is $c_{ij}^{\,2} = \sum_{k=1}^{n} c_{ik} \cdot c_{kj}$. The two indirect links from node i to j are represented by $c_{ik} \cdot c_{kj}$ (Taaffe & Gauthier p. 122). The original graph is show in the top left corner of the graphic. The resultant matrix, C^2 is show in the bottom right corner of the graphic. The presence or absence of two-link paths can be determined by the resultant matrix C^2. The two-link paths are represented by non-zero entries in the matrix. The presence of a zero in the matrix indicates that neither a direct or indirect path of only two links exists.



C

Nodes	V_1	V_2	V_3	V_4	V_5	V
V_1	0	1	0	0	0	0
V_2	1	0	0	0	1	0
V_3	0	0	0	1	0	0
V_4	0	0	1	0	1	0
V_5	0	1	0	1	0	1
V_6	0	0	0	0	1	0

V_1	V_2	V_3	V_4	V_5	V_6
0	1	0	0	0	0
1	0	0	0	1	0
0	0	0	1	0	0
0	0	1	0	1	0
0	1	0	1	0	1
0	0	0	0	1	0

C^2

Nodes	V_1	V_2	V_3	V_4	V_5	V_6
V_1	1	0	0	0	1	0
V_2	0	2	0	1	0	1
V_3	0	0	1	0	1	0
V_4	0	1	0	2	0	1
V_5	1	0	1	0	3	0
V_6	0	1	0	1	0	1

Figure 3-6 Matrix Multiplication

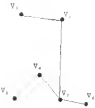
The resultant matrix, C^2, tells us which pairs of nodes have two-link paths connecting them. For example, node 1 and 5 contain a two link path between them. If we look at the original matrix (C), we find that there is no direct connection between these two nodes. Node 3 and 4 contain a direct link between them, but not a two-link path. Note that the most distant nodes in the network are 1 and 3.

If the new matrix (C^2) is multiplied by the original matrix (C), the number of three link paths will be identified in the product matrix. Figure 3-7shows the matrices used to determine the three-link paths in the network, as well as the original network. The new matrix introduced, C^3, provides the connection by three-link paths between nodes. The new matrix was produced by multiplying the rows of the original matrix (C) by the columns of the second matrix (C^2). More pairs of nodes in the network are connected by third-link paths than are connected by either two-link paths or direct paths. By using this same procedure of multiplication any number of link paths can be determined.

By adding the number of paths between nodes on the network, direct, two link, and three link, the accessibility matrix or T, can be created. Figure 3-7 illustrates each matrix added to create the accessibility matrix (T). The resulting accessibility matrix is shown in Figure 3-7. By summing each row in the accessibility matrix, we are able to determine the column vector, which is the accessibility of a given node in a network. Figure 3-7 illustrates the addition of rows and resultant vectors. The summation of elements: $\sum_{k=1}^{d} c_{ij}^{\ k}$ for k=1...d matrices. The summation of typical elements of c_{ij} of the matrices up to the power d will yield the typical elements of the accessibility matrix (A) with typical elements a_{ij} (see Figure 3-7). A diameter of four represents the number of links between the two most distant nodes, 1 and 3 are separated by four links.

If the were then ranked based on the accessibility vectors, node 5 would be the most accessible, as it has the highest vector. By summing the rows across the accessibility matrix, row 5 produced the highest value. This means that node five is the most connected and accessible node in the network. Nodes 4 and 2 have the second highest value, followed by node 6. Finally, the least two accessible nodes in the network would be 1 and 3; their row totals produced the lowest values.

Though it's possible to continue multiplying the matrices, the network cannot have more than v-1 paths without establishing redundant paths. Multiplication is usually



C

Nodes	V1	V2	V3	V4	V5	V6
V1	0	1	0	0	0	0
V2	1	0	0	0	1	0
V3	0	0	0	1	0	0
V4	0	0	1	0	1	0
V5	0	1	0	1	0	1
V6	0	0	0	0	1	0

C^3

Nodes	V1	V2	V3	V4	V5	V6
V1	0	2	0	1	0	1
V2	2	0	1	0	4	0
V3	0	1	0	2	0	1
V4	1	0	2	0	4	0
V5	0	4	0	4	0	3
V6	1	0	1	0	3	0

C^2

Nodes	V1	V2	V3	V4	V5	V6
V1	1	0	0	0	1	0
V2	0	2	0	1	0	1
V3	0	0	1	0	1	0
V4	0	1	0	2	0	1
V5	1	0	1	0	3	0
V6	0	1	0	1	0	1

C^4

Nodes	V1	V2	V3	V4	V5	V6			
V1	2	0	1	0	4	0	Σ	=	7
V2	0	6	0	5	0	4	Σ	=	15
V3	1	0	2	0	4	0	Σ	=	7
V4	0	5	0	6	0	4	Σ	=	15
V5	4	0	4	0	11	0	Σ	=	19
V6	0	4	0	4	0	3	Σ	=	11
	Network Connectivity							=	74

Accessibility Matrix (A)

Nodes	V1	V2	V3	V4	V5	V6			
V1	3	3	1	1	5	1	Σ	=	14
V2	3	8	1	6	5	5	Σ	=	28
V3	1	1	3	3	5	1	Σ	=	14
V4	1	6	3	8	5	5	Σ	=	28
V5	5	5	5	5	14	4	Σ	=	38
V6	1	5	1	5	4	4	Σ	=	20
	Network Accessibility							=	142

Figure 3-7. Three-linkage paths and the accessibility matrix
Note: Accessibility matrix (A) is a composite of C, C^2, C^3, and C^4, based on the sum of the typical elements of these matrices

stopped when the network has reached its diameter, because each of the nodes in the network is connected to each of the other nodes in the network. The matrix is powered to it's diameter, in this case d=4, to ensure that the summation of for all matrices up to that power will yield non-zero entries in the resultant matrix. The greatest number of links between any two nodes is four. This does not mean that there will be no redundancies in the network Powering up to the diameter ensures that, all of the nodes are connected to each other node in the network by some path, be it direct or indirect (via multiple links or paths).

These techniques first appeared in a study on the Interstate Highway System in the United States (Garrison 1960). Garrison's study is noted by Taaffe and Gauthier in their 1973 publication the Geography of Transportation. Garrison, as well as Taaffe and Gauthier raise concern over the redundancies that arise using matrix multiplication, as they many not be any representation of distance minimizing behavior of networks. In response to these concerns, a shortest-path matrix multiplication procedure was discussed. This procedure was introduced by Alfonso Shimbel in 1953. Shimbel's procedure computes accessibility in terms of the distance between nodes, thereby eliminating redundancies. The procedure calculates the length of the shortest path between nodes Taaffe and Gauthier (1973) suggest that for many real-world problems redundancies are of no importance (p. 132) For this particular research however, the importance of redundancy in long-haul fiber networks is pertinent. The majority of the links in this network are redundant, which must be considered in the practical application of link removals due to disturbance. These redundancies increase the insurance of a network's functionality should it experience a disturbance. They allow also allow for increased traffic and data transfer during normal conditions.

To review, the connectivity of a node is determined by summing across the rows of the resultant matrix, and by summing the rows of the resultant matrix, the connectivity

of the entire network can be determined (see the, Matrix C^4 Network Connectivity in Figure 3-7). The measures of accessibility and connectivity give nearly identical measures in nodal/link importance to a network (compare in Figure 3-7). Both accessibility and connectivity measures have been demonstrated in this chapter. However, the following analysis concentrates on the connectivity of nodes in order to determine which nodes are best connected and most important to the network rather than which nodes are most accessible.

Weighted or Valued Graphs and Shortest-Path

Networks can be represented as "valued graphs." This adds weights to the network's links, highlighting their capacity or flow potential, or distance. If we take the original graph and assign weights to the links we have a weighted network. Figure 3-8 illustrates the network as a weighted graph, with values assigned to each link. The new weighted network is represented in a matrix with the new weighted values in Table 3-8. The values added to the weighted network represent distance between the corresponding pairs of nodes. In Table 3-8, the length (or value) of existing direct connections between pairs of nodes is represented in the corresponding matrix cells. For those pairs of nodes that do not have direct connections between them, a ∞ is recorded in the cell, representing infinity. Connection between a node and itself is meaningless, so those cells have a zero recorded in the cell. These self-connection cells always fall along the main diagonal of the matrix. We can tell from the matrix the distance between nodes in direct connections, and also indirect connections. To travel from node 1 to node 2 a distance of 20 is traveled, accomplished in a direct connection. To travel from node 3 to 5 however, an indirect connection involving node 4 is required.

The distance from node 3 to 4 is 20, and the distance from node 4 to 5 is 10. The total distance from node 3 to 5 is then 30.

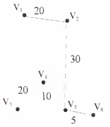


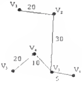
Figure 3-8. Weighted network

Table 3-8. Weighted network represented as a matrix

Nodes	V_1	V_2	V_3	V_4	V_5	V_6
V_1	0	20	∞	∞	∞	∞
V_2	20	0	∞	∞	30	∞
V_3	∞	∞	0	20	∞	∞
V_4	∞	∞	20	0	10	∞
V_5	∞	30	∞	10	0	5
V_6	∞	∞	∞	∞	5	0

In order to answer questions of accessibility questions about more complex matrices, a procedure similar to matrix multiplication is required. The new procedure is less complex than matrix multiplication. Element-by-element addition $(x\ y = x + y)$ replaces element-by-element multiplication, and instead of summing the values the minimum value is inserted in the appropriate cell of the new matrix $[x + y = \min(x\ y)]$. The value *ij* becomes the minimum value of the sums of these two-stage links from origin *i* to *k* and then to destination *j*, or $c_{ij}^{2} = \sum_{k=1}^{n} c_{ik} \cdot c_{kj} = \min(c_{ik} + c_{jk})$. The previous equation for summing the products in matrix multiplication was : $C_{ij}^{2} = \sum_{k=1}^{n} c_{ik} \cdot c_{kj}$ Figure 3-1 illustrates the element-by-element addition and determination of the minimum sum. It also indicates where the new value would be inserted in the new matrix, L^2. In the example shown in Figure 3-9, the least-path linkage between nodes 1 and 5 are determined.

The two-step path that connects node 1 and 5 has a length of 50. By using this methodology, the successive powers of the weighted matrix are calculated. The results indicate the minimum distance required between each pair of nodes. When a value of *m* has been reached, a matrix of minimum distance has been achieved. Once the matrix contains no non-zero entries, with the exception of the main diagonal, minimum distance has been achieved between each and every pair of nodes in the network (Taaffe & Gauthier, p. 141). This method has been compared to Shimbel's minimum connectivity procedure, which includes powering the matrix until there are no zero connections left in the matrix (Shimbel 1953, Taffe & Gauthier 1973). Taffe and Gauthier compared the addition method with Shimbel's method and found that the structural relationships don't change but the distance criteria might give a more refined measurement of nodal accessibility.



Nodes	V_1	V_2	V_3	V_4	V_5	V_6
V_1	0	20	∞	∞	∞	∞
V_2	20	0	∞	∞	30	∞
V_3	∞	∞	0	20	∞	∞
V_4	∞	∞	20	0	10	∞
V_5	∞	30	∞	10	0	5
V_6	∞	∞	∞	∞	5	0

Nodes	V_1	V_2	V_3	V_4	V_5	V_6
V_1	0	20	∞	∞	∞	∞
V_2	20	0	∞	∞	30	∞
V_3	∞	∞	0	20	∞	∞
V_4	∞	∞	20	0	10	∞
V_5	∞	30	∞	10	0	5
V_6	∞	∞	∞	∞	5	0

$$
\begin{aligned}
(V_1 \rightarrow V_1) + (V_1 \rightarrow V_5) &= 0 + \infty &&= \infty \\
(V_1 \rightarrow V_2) + (V_2 \rightarrow V_5) &= 20 + 30 &&= 50 \\
(V_1 \rightarrow V_3) + (V_3 \rightarrow V_5) &= \infty + \infty &&= \infty \\
(V_1 \rightarrow V_4) + (V_4 \rightarrow V_5) &= \infty + 10 &&= \infty \\
(V_1 \rightarrow V_5) + (V_5 \rightarrow V_5) &= \infty + 0 &&= \infty \\
(V_1 \rightarrow V_6) + (V_6 \rightarrow V_5) &= \infty + 5 &&= \infty
\end{aligned}
$$

Nodes	V_1	V_2	V_3	V_4	V_5	V_6
V_1					50	
V_2						
V_3						
V_4						
V_5						
V_6						

Figure 3-9. Indirect Connections in a Weighted Network

Node and Link Removal

To determine the degree of importance of any node to a network, it can be removed from the network (Haggett, Cliff, & Frey 1977b, p. 322). The remaining nodes in

the network are then recalculated to determine changes in the network; diameter (r), row totals (R_i), network connectivity; $\sum_{i=1}^{n} R_i^{(r)}$, and disconnects. The matrix is repowered after a node or pair of nodes has been removed. The results are compared to the original network values. When a network has been powered to it's diameter ($c_{ij} \neq 0$), the row sums of the resultant matrix indicate the connectivity of each node in the network.

By summing the row totals, the connectivity of the entire network can be derived. The network connectivity index (NCI) is obtained by summing the powered matrix's row totals; $\sum_{i=1}^{n} R_i^{(r)}$. Nodes which cause the largest changes in network connectivity, to i and j values, will exercise the greatest degree of importance within the network (Haggett, Cliff, & Frey 1977b).

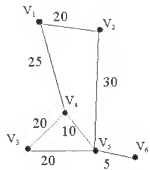
The method is demonstrated using the network shown in Figure 3-10. The network is comprised of six nodes and seven links. The network was powered to a diameter of three, and the resultant matrix is shown (Figure 3-10). The NCI of the complete network is 34525000 (the sum of row totals). Figure 3-11 illustrates the impact of node removal upon a network. Node V3 was removed from the network. This removal reduces the network to five nodes and five links (Figure 3-11). This network was powered to a diameter of three. The sum of rows of the resultant matrix was less than that of the complete network, as expected. With the removal of Node V3, the connectivity of the network dropped to 14625000. This is a 57.6% decrease in the NCI value. Figure 3-12 illustrates the removal of a link from a network The link connecting node V4 and node V5 was removed from the network The remaining network was powered to a diameter of three. The row totals of the resultant matrix were summed to determine the connectivity of the network with absence of this link. The connectivity drops to 24205000, decreasing the connectivity of the network 30%. The NCI % change is relatively high for the examples illustrated in Figures 3-11 and 3-12, however the network contains a small number of both links and nodes. The node removal reduced

the number of nodes in the network 16.6% (from 6 nodes to 5),while the link removal reduced the number of connections in the network 14% (from 7 links to 6).

Conclusion

The methods discussed in this chapter can be used to assess changes in connectivity and accessibility, at node and network levels, in response to the removal of a link or node. This chapter has discussed data methodology and analysis that are used in the following chapters. The U.S. Internet backbone network is analyzed in Chapters 4 and 5. Node removal and matrix multiplication have been introduced in this chapter, and they are they main tools used in the analysis for both Chapter 4 and 5. Network changes will be measured for each removal scenario by comparing measurements in diameter, row totals, NCI, and ranking, to the complete network. Chapter 4 describes the unweighted analysis of the U.S. Internet backbone network. Single and pair node removal scenarios performed to determine the effects upon the network. The analysis performed in Chapter 5 also uses matrix multiplication and node and link removal techniques, adding weights to the network. A better representation of the actual Internet backbone network is achieved by using valued links.

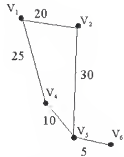
Nodes	V1	V2	V3	V4	V5	V6
V1	0	20	∞	25	∞	∞
V2	20	0	∞	∞	30	∞
V3	∞	∞	0	20	20	∞
V4	25	∞	20	0	10	∞
V5	∞	30	20	10	0	5
V6	∞	∞	∞	∞	5	0



Nodes	V1	V2	V3	V4	V5	V6	Row Totals		
V1	2023125	300000	1082500	440000	2182500	50000	Σ	=	6078125
V2	300000	2712500	1435000	2067500	440000	298750	Σ	=	7253750
V3	1082500	1435000	1340000	950000	950000	182500	Σ	=	5940000
V4	440000	2067500	950000	2108125	1060000	197500	Σ	=	6823125
V5	2182500	440000	950000	1060000	2953125	40000	Σ	=	7625625
V6	50000	298750	182500	197500	40000	35625	Σ	=	804375
				Network Connectivity Index (NCI)			**Σ**	**=**	**34525000**

Figure 3-10 Example of a complete network

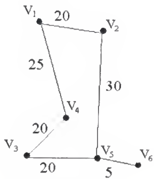
Nodes	V1	V2	V4	V5	V6
V1	0	20	25	∞	∞
V2	20	0	∞	30	∞
V4	25	∞	0	10	∞
V5	∞	30	10	0	5
V6	∞	∞	∞	5	0



Nodes	V1	V2	V4	V5	V6		
V1	1773125	0	0	1742500	0	Σ =	3515625
V2	0	2352500	1627500	0	238750	Σ =	4218750
V4	0	1627500	1168125	0	157500	Σ =	2953125
V5	1742500	0	0	1773125	0	Σ =	3515625
V6	0	238750	157500	0	25625	Σ =	421875
					Network Connectivity Index (NCI)	Σ =	146250000

Figure 3-11 Example of a Node removed from a network

Nodes	V1	V2	V3	V4	V5	V6
V1	0	20	∞	25	∞	∞
V2	20	0	∞	∞	30	∞
V3	∞	∞	0	20	20	∞
V4	25	∞	20	0	∞	∞
V5	∞	30	20	∞	0	5
V6	∞	∞	∞	∞	5	0



Nodes	V1	V2	V3	V4	V5	V6			
V1	1660625	300000	912500	240000	1410000	50000	Σ	=	4573125
V2	300000	2322500	1275000	1162500	200000	258750	Σ	=	5518750
V3	912500	1275000	1260000	300000	300000	172500	Σ	=	4220000
V4	240000	1162500	300000	1460625	940000	75000	Σ	=	4178125
V5	1410000	200000	300000	940000	2275625	0	Σ	=	5125625
V6	50000	258750	172500	75000	0	33125	Σ	=	589375
				Network Connectivity Index (NCI)			**Σ**	**=**	**24205000**

Figure 3-12 Example of a link removed from a network

CHAPTER 4
THE U.S. INTERNET BACKBONE NETWORK: AN UNWEIGHTED ANALYSIS

Protection of critical infrastructure in the United States has been a hot topic in recent months. One of the most pressing issues is the insurance and protection of critical infrastructure. The protection of infrastructures that directly effect the economy and financial sector is needed. Major disturbances to the Internet's infrastructure have occurred, which directly effects, among other things, the financial market. The fall of the Twin Towers during the terrorist attacks of 9/11 and the northeastern power outages of August 2003 have each led to Internet disruption due to physical failure.

It must be noted that disturbances to physical Internet infrastructure occur frequently on small scales. These disturbances are usually accidental, often caused by backhoes and shovels. These disturbances are less publicized, in part because the disruption is minor and generally effects local service.

September 11th, 2001 was the largest loss of physical telecommunication infrastructure (FCC 2001). Verizon alone lost their central office[1] along with 182,000 voice circuits, more than 1.6 million data circuits, and more than 11,000 lines serving Internet service providers (GAO 2003). With the loss of Verizon's central office, 34,000 businesses lost their telecommunication service. Verizon is an example of just one telecommunication provider that was dramatically effected by the terrorist attacks, many other telecommunication providers lost physical infrastructure. Though the exchange and clearing organizations were undamaged by the terrorist attacks, the economic disruption was severe due to the loss of telecommunication infrastructure (GAO 2003).

[1]A central office is a facility, owned & operated by a telecommunication firm, which houses the switching equipment that links customers to voice and data networks within and outside the service area.

A power blackout on August 14, 2003, simultaneously affected Detroit, Cleveland. Columbus and Long Island, New York as well as Canadian cities, Toronto, and Ottowa (Rosenblum 2003, Semple 2003, NASA 2003). Figures 4-1 and 4-2, courtesy of Chris Elvidge of the U.S. Air Force, have been made available by the NASA Earth Observatory in their collection of unique images. Figures 4-1 and 4-2 illustrate the widespread power outage. The top image was taken on August 14, 2003, roughly twenty hours before the blackout occurred. The lower image was taken on August 15, 2003, about seven hours after the blackout began: a post-September 11th reminder to the vulnerability of our infrastructure and the interconnectedness of our cities. A series of events caused a domino effect across interconnected power grids. The result was a widespread outage (NASA 2003). The cities affected by the blackout experienced varying degrees of impact upon themselves and their neighbors. The same type of domino effect has occurred within the nation's telecommunication infrastructure network. Telecommunication and computer equipment is often dependent upon the power grid, though there are exceptions of power generation to avoid the power grid dependency (particularly in data and colocation centers [interconnection hubs for networks]).

Sixty Hudson Street, located in the financial district in Manhattan, is one of the main hubs of Internet activity and interconnection in the world. Unlike other data and interconnection facilities, this building is reliant upon the power grid and weathered major disturbance during the power outages in August of 2003. With little time bought by generators, many companies that housed equipment in the building were soon negatively affected (Careless 2003). Figure 4-1 further illustrates the interdependencies between Internet infrastructure and the power grid. Figure-4-3 shows a map of Internet routing outages that occurred during the August 14, 2003 power outages. The interdependency of networks increases their vulnerability to disturbances, as illustrated by the terrorist attacks of 9-11, the power outages and other disturbances.

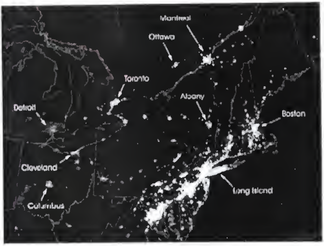


Figure 4-1. Image taken August 14, 2003, 9:29 p.m. EDT, about 20 hours before blackout (NASA Earth Observatory, 2003)

Figure 4-2 Image taken August 15, 2003, 9:14 p.m. EDT, about 7 hours after blackout (NASA Earth Observatory 2003)

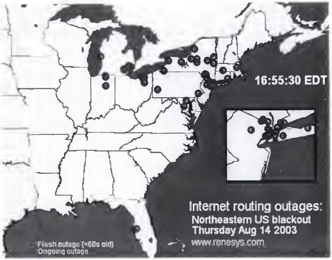


Figure 4-3. Internet routing outages during the 2003 blackout. Source: www renesy.com

This Chapter focuses on long-haul fiber, where, like the power grids across the U.S., infinite overlaps and interconnection occur to create one large network. The disturbance of each node holds the potential to disrupt the overall network. The important a node is to the overall network, the more effect the disturbance will have upon the connectivity of the network. A disruption in the network can also have major effect on sectors reliant upon that network. Economic sectors are increasingly dependent upon telecommunication infrastructure, particularly long-haul fiber. Should particular long-haul fiber routes endure a major disturbance, the economy would be directly affected. Detecting vulnerabilities is the beginning of protection.

This chapter introduces a methodology to identify the most critical links and nodes in the Internet Backbone Network U.S. The analysis will proceed in two stages. First, an unweighted analysis will consider each link to be of equal importance within the network. In short, the unweighted analysis will not take into account the amount of bandwidth that

connects a node to other nodes. Weights will be added in Chapter 5, recognizing the amount of bandwidth each link represents. In this chapter, a method for identifying a hierarchy of nodes and links within an undirected, unweighted network is introduced. The method applies an unweighted scenario to the Internet Backbone Network in the United States using the graph theoretic concepts discussed in Chapter 3; this chapter will assess the vulnerability of the long-haul fiber network in the U.S. on several different levels:

- How node removal/disruption affects nodality (and ranking) of all other nodes and changes the ranking of all remaining links in the network
- How node removal/disruption affects the overall connectivity of the network as measured by the sum total of all nodality indices after removal/disruption;
- How link removal/disruption affects the nodality (and ranking) of nodes
- How link removal/disruption affects the overall connectivity of the network as measured by the sum total of all nodality indices
- Changes in the rank and importance of remaining links.

A total of 218 nodes representing cities in the U.S. (including Alaska and Hawaii) were used for the unweighted and weighted analysis (Figure 4-4). Table 4-1 shows a complete list of the C/MSAs used in the analysis. Basic network measurements were performed using the Internet Backbone Network data to determine network connectivity. A symmetrical binary connectivity matrix is used to represent the network. The number of binary links represented in the matrix (1042) were divided in half (521), to identify the total number of unique links in the network (to avoid double counts). Figure 4-5 illustrates the network links used in both the unweighted and weighted analysis. This partially addressed the redundancy that occurs with matrix representation of a network. The self-to-self links that are represented in the principal diagonal will not be included in the totals. As it is assumed that nodes are not connected to themselves. The gamma and alpha indices: $\gamma = \frac{actual\ edges}{\max edges} = \frac{e}{e\max} = \frac{e}{3(v-2)} = \frac{2v-3}{3(v-2)}$ and $\alpha = \frac{actual\ circuits}{\max circuits} = \frac{e-v+1}{2v-5} = \frac{v-2}{2v-5}$.

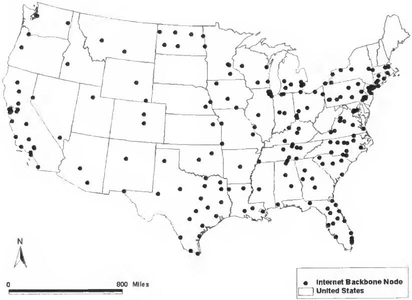


Figure 4-4. Nodes of U.S. Internet backbone network
Note: Honolulu, HI & Anchorage, AK are not included in this map, but are Internet Backbone Nodes used in this analysis

Figure 4-5. Links of U.S. Internet backbone network

Table 4-1. List of consolidated metropolitan statistical areas (CMSA) and metropolitan statistical areas (MSA) in U.S. Internet backbone network

C/MSA	C/MSA
Abilene	Charlottesville
Albany-Schenectady-Troy	Chattanooga
Albuquerque	Cheyenne
Alexandria	Chicago-Gary-Kenosha
Allentown-Bethlehem-Easton	Chico-Paradise
Altoona	Cincinnati-Hamilton
Amarillo	Clarkrange
Anchorage	Cleveland-Akron
Ardmore	Colorado Springs
Atlanta	Columbia
Austin-San Marcos	Columbus
Bakersfield	Cookeville
Baton Rouge	Corpus Christi
Bellevue	Dallas-Fort Worth
Billings	Danville
Birdstown	Daytona Beach
Birmingham	Dayton-Springfield
Bismarck	Denver-Boulder-Greeley
Blacksburg	Des Moines
Bohemia	Detroit-Ann Arbor-Flint
Boise	Devils Lake
Boonshill	Dibrell
Boston-Worcester-Lawrence	Dickinson
Bowling Green	Dunlap
Bridgeport	Durand
Brownsville-Harlingen-San Benito	Eau Claire
Bryan-College Station	El Paso
Buffalo-Niagara Falls	Elkhart-Goshen
Casper	Emeryville
Cedar Knolls	Erie
Celina	Estill Springs
Chapel Hill	Eugene-Springfield
Charleston-North Charleston	Fargo-Moorhead
Charlotte	Fayetteville
Charlotte-Gastonia-Rock Hill	Flat Creek
Florence	Kansas City
Fort Myers-Cape Coral	Lafayette
Fosterville	Lakeland-Winterhaven
Freehold	Lancaster
Fresno	Laredo
Gainesboro	Las Vegas
Gainesville	Laurinburg
Garden City	Lexington-Fayette
Gardena	Lincoln
Glenview	Little Rock
Grand Forks	Livingston
Grand Rapids-Muskegon-Hollan	Longview-Marshall
Green Bay	Los Angeles-Riverside-Orange County

Table 4-1. Continued

C/MSA	C/MSA
Greensboro-Winston Salem-High Point	Louisville
Greenvilee-Spartanburg-Anderson	Lumberton
Greenwood	Macon
Hackensack	Madison
Hamilton Square	Market Place
Harrisburg-Lebanon-Carlisle	McAllen-Edinburg-Mission
Harrisonburg	Medford-Ashland
Hartford	Melbourne-Titusville-Palm Bay
Hayward	Memphis
Helena	Miami-Fort Lauderdale
Hollywood	Milwaukee-Racine
Honolulu	Minneapolis-St. Paul
Houston-Galveston-Brazoria	Minot
Hudson	Mobile
Huntsville	Monroe
Indianapolis	Montgomery
Jackson	Nashville
Jacksonville	New Bern
Jamestown	New Brunswick
Johnstown	New London-Norwich
Joplin	New Market
Kalamazoo-Battle Creek	New Orleans
New York-Northern New Jersey-Long Island	Saint Augustine
Norfolk-Virginia Beach-Newport News	Salinas
Oakbrook	Salt Lake City-Ogden
Ocala	San Antonio
Odessa-Midland	San Diego
Ojus	San Francisco-Oakland-San Jose
Oklahoma City	San Luis Obispo-Atascadero-Paso Robles
Omaha	Santa Barbara-Santa Mria-Lompoc
Orlando	Scranton-Wilkes Barre-Hazleton
Owantonna	Seattle-Tacoma-Bremerton
Palmdale	Sherman Oaks
Pennsauken	Shreveport-Bossier City
Pevley	South Bend
Philadelphia-Wilmington-Atlantic City	South Holland
Phoenix-Mesa	Southfield
Pittsburgh	Spencer
Pittsfield	Spokane
Plevna	Springfield
Plymoth	St. Louis
Pompano	Staunton
Portland-Salem	Stockton-Lodi
Providence-Fall River-Warwick	Syracuse
Raliegh-Durham-Chapel Hill	Tallahassee
Redding	Tampa-St.Petersburg-Clearwater
Redwood City	Toledo

Table 4-1. Continued

C/MSA	C/MSA
Reno	Topeka
Richmond-Petersburg	Tracy City
Roachdale	Tulsa
Roanoke	Tuscan
Roanoke Rapids	Victoria
Rochelle Park	Waco
Rochester	Wartburg
Rocky Mount	Washington-Baltimore
Rolling Meadows	Waterford
Sacramento-Yolo	Wayne
West Haven	
West Palm Beach-Boca Raton	
Williamsburg	
Willimsport	
Williston	
Wilmington	
Woodberry	
Youngstown-Warren	

The gamma and alpha indices were computed as standard measures of network connectivity. The long-haul fiber network contains 218 nodes or vertices and 521 links or edges. The gamma index is simply the ratio of the number of nodes in a network to the maximum number possible in that network: $\gamma = \frac{e}{3(v-2)} = \frac{521}{648}$. Hence, the gamma index for the Internet backbone network in the U.S. is 0.804. In terms of maximal connectivity, this means that the network is 80.4% connected.

The alpha index, is a ratio measure of the number of actual links to the maximum number possible in the network: $\alpha = \frac{actual\ circuits}{\max circuits} = \frac{e-v+1}{2v-5} = \frac{521-218+1}{2(218)-5} = \frac{304}{431} = 0.705$. Like the gamma index, the alpha index ranges from 0-1. A zero value would represent a network that is minimally connected, maximally connected networks would be represented by 1. The network linkage, or circuitry, is 70.5 % of the maximum.

Table 4-2 shows binary data for the Internet backbone network in the U.S. for 2003. The binary totals represent the number of nodes to which a node is directly linked. Chicago-Gary-Kenosha has 36 binary links, more than any other metropolitan area. This means that Chicago was directly connected to 36 other cities within the U.S. Internet

backbone network. New York and Dallas followed closely with 33 and 30 links respectively. The 2003 data contains data for 218 metropolitan areas in the U.S. As shown in Table 4-2 Washington-Baltimore has 28 separate long-haul links to metropolitan areas, ranking fourth. Atlanta and San Francisco-Oakland-San Jose have 27 and 26 binary links, respectively, filling the fifth and sixth ranks. These cities have the most binary connections to other major metropolitan areas in the U.S.

Table 4-2. Long-haul fiber optic binary connections in U.S. metropolitan areas

Metropolitan Area (C/MSA)	Binary Connections 2003 (Rank)
Chicago-Gary-Kenosha	36 (1)
New York-Northern New Jersey-Long Island	33 (2)
Dallas-Fort Worth	30 (3)
Washington-Baltimore	28 (4)
Atlanta	27 (5)
San Francisco-Oakland-San Jose	26 (6)
Denver-Boulder-Greeley	19 (7)
Los Angeles-Riverside-Orange County	19 (7)
Cleveland-Akron	18 (9)
Kansas City	18 (9)
Sacramento-Yolo	16 (11)
Houston-Galveston-Brazoria	15 (12)
Miami-Fort Lauderdale	14 (13)
St. Louis	14 (13)
Boston-Worcester-Lawrence	13 (15)
Nashville	13 (15)
Seattle-Tacoma-Bremerton	13 (15)
Tampa-St.Petersburg-Clearwater	13 (15)

Matrix multiplication was the main tool used in the unweighted analysis. As explained and demonstrated in Chapter 3, matrix multiplication can help to determine nodal accessibility as well as to establish a ranking of nodes and links within a network. The connectivity matrix is a model of the network, allowing various scenarios to be simulated in order to learn more about the network. By using matrix multiplication, the most critical nodes to the overall network were determined. Various link-removal scenarios were carried out using matrix multiplication. Both single-node and pairs-of-nodes removal scenarios were performed. Once the city ranking for the complete

network was established, the nodes to be used in the single-and-paired-node-removal scenarios were identified. Each of the nodes that were ranked in the top 12 in terms of connectivity for the original, fully connected network were used in the single-node removal scenarios. Every possible pairing of the top 12 nodes was also removed from the network, accounting for 72 double-node removal scenarios. The new network measurements show the degree of a node's importance to the network, and the effect of different removal scenarios upon the entire network: diameter, disconnects, the Unweighted Relative Connectivity Index (URCI), the Network Connectivity Index (NCI) and the percentage change in connectivity.

For each removal scenario, the binary connectivity matrix was multiplied to its diameter. This means multiplying the entire matrix until no zeros exist within the matrix. In matrix multiplication, a zero represents a disconnect. The absence of zeros illustrates each node in the network is connected to each of the other nodes in the network. When a network reaches its diameter, all of the nodes in the network are connected. The complete matrix was multiplied to its diameter, 11. This means the network was multiplied 10 times before each node in the network was connected to each of the other nodes in the network through some path.

Various removal scenarios completely disconnected particular nodes from the entire network. These are called "disconnects." A disconnected node is severed from the matrix and connection will not occur by continued powering of the matrix. This was considered when determining the presence or absence of zero's within the connected nodes of the network.

After the matrix was powered to its diameter, the rows from the product of the last powering were totaled. These totals were used to determine which 25 nodes were at the top of the hierarchy (Table 4-3). Figure 4-6 illustrates the hierarchy of the 218 nodes in the U.S. based on the unweighted relative connectivity index. An index was then created

from the row totals. Each row total was divided by the minimum row total of the top 25 nodes in the network. The minimum observation was Salt Lake City, with a row total of 538479367663. This index is an unweighted relative connectivity index (URCI). The value for Salt Lake City-Ogden was reassigned to 1.0 to begin the index.

$URCI_i = \frac{x_i}{x_{min}}$ $i = 1..N$, where N=25. For example, to calculate the URCI values for Atlanta (row totals= 1544357876452) within a fully connected network,

$URCI_i = \frac{x_i}{x_{min}} = \frac{1544357876452}{538479367663} = 2.868$. Figure 4-7 shows the geographical distribution of the top 25 nodes based on the URCI.

The network connectivity index (NCI) was obtained by summing the row totals for the entire network. The NCI was created after the row totals for each multiplied matrix scenario had been converted to the URCI. The converted values were summed to give a value representing the overall connectivity for the entire network for each removal scenario. The simple NCI model is: $x_{total} = \sum x_i$ where i = 1...218. The NCI of the fully connected network is 99.499. This means that the summed row totals, after being converted using the URCI, had a value of 99.499.

The percentage change value is used to illustrate the amount of change in network connectivity from each network scenario compared to the original, fully connected network. The change was calculated by subtracting the NCI (of a given scenario) from the original connectivity value. The result was then divided by the original value, 99.499. The equation is $\%\Delta\chi = \frac{(x_1 - x_2)}{x_1}$. For example, to calculate the percentage change in the connectivity of the network with Chicago-Gary-Kenosha (connectivity 36.909) removed in comparison to the original network the equation is

$\%\Delta\chi = \frac{(99.499 - 36.909)}{99.499} = 0.692$ The final result gives the percentage of change in network connectivity compared to the fully connected network.

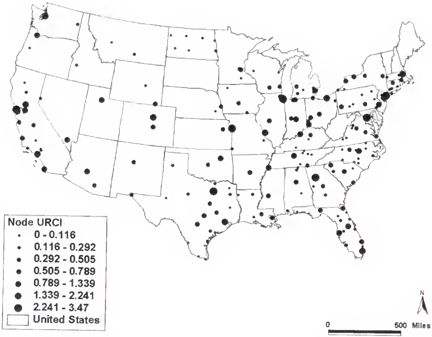


Figure 4-6. Cities of U.S. Internet backbone network based on URCI, using graduated symbols to represent connectivity importance.

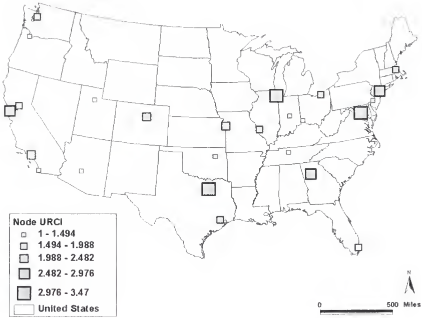


Figure 4-7. Top 25 nodes in U.S. Internet backbone network, using graduated symbols based on URCI bullets.

Table 4-3. U S. city ranking for the Internet backbone network

Rank	C/MSA	URCI
1	Chicago-Gary-Kenosha	3.470
2	Washington-Baltimore	3.217
3	Dallas-Fort Worth	3.162
4	San Francisco-Oakland-San Jose	2.957
5	Atlanta	2.868
6	New York-Northern New Jersey-Long Island	2.797
7	Kansas City	2.467
8	Denver-Boulder-Greeley	2.241
9	Los Angeles-Riverside-Orange County	2.239
10	St. Louis	1.901
11	Sacramento-Yolo	1 878
12	Seattle-Tacoma-Bremerton	1.650
13	Boston-Worcester-Lawrence	1.601
14	Houston-Galveston-Brazoria	1.595
15	Cleveland-Akron	1.585
16	Miami-Fort Lauderdale	1.517
17	Indianapolis	1.339
18	San Diego	1.293
19	Phoenix-Mesa	1.259
20	Pennsauken	1.229
21	Tulsa	1.218
22	Nashville	1.156
23	Cincinnati-Hamilton	1.054
24	Portland-Salem	1.025
25	Salt Lake City-Ogden	1.000

Note: NCI 99.499; diameter 11; fully connected

Complete Matrix

Node Ranking

The ranking of nodes for the complete, unweighted network shows little surprise in terms of the top five cities. Here, Chicago-Gary-Kenosha, Washington-Baltimore, Dallas-Fort Worth, San Francisco-Oakland-San Jose and Atlanta comprise the top five nodes in terms of connectivity in the unweighted network. Table 4-4 shows the ranking of nodes for the complete network and the number of redundant connections a node contains, as well as the ranking of the city based on number of redundant connections (Table 4-4).

In general, the URCI ranking corresponds relatively closely to redundant connection ranking (Table 4-4). The ranking does pose some interesting shifts in comparison to ranking based on other types of Internet infrastructure (Table 4-5).

Table 4-4. U.S. city ranking for Internet backbones and redundant connections

Rank	MSA	URCI	Redundant Connections 2003 (Rank)
1	Chicago-Gary-Kenosha	3.470	141 (5)
2	Washington-Baltimore	3.217	245 (1)
3	Dallas-Fort Worth	3.162	142 (4)
4	San Francisco-Oakland-San Jose	2.957	147 (3)
5	Atlanta	2.868	105 (6)
6	New York-Northern New Jersey-Long Island	2.797	183 (2)
7	Kansas City	2.467	59 (11)
8	Denver-Boulder-Greeley	2.241	65 (10)
9	Los Angeles-Riverside-Orange County	2.239	38 (7)
10	St. Louis	1.901	38 (20)
11	Sacramento-Yolo	1.878	46 (18)
12	Seattle-Tacoma-Bremerton	1.650	58 (13)
13	Boston-Worcester-Lawrence	1.601	59 (12)
14	Houston-Galveston-Brazoria	1.595	69 (8)
15	Cleveland-Akron	1.585	64 (10)
16	Miami-Fort Lauderdale	1.517	55 (15)
17	Indianapolis	1.339	26 (30)
18	San Diego	1.293	32 (25)
19	Phoenix-Mesa	1.259	37 (23)
20	Pennsauken	1.229	14 (43)
21	Tulsa	1.218	20 (37)
22	Nashville	1.156	24 (31)
23	Cincinnati-Hamilton	1.054	11 (55)
24	Portland-Salem	1.025	38 (21)
25	Salt Lake City-Ogden	1.000	28 (28)

The same cities are constantly jostling for top rank: Chicago-Gary-Kenosha, Washington-Baltimore, Dallas-Fort Worth, San Francisco-Oakland-San Jose, New York-Northern New Jersey-Long Island, and Los Angeles-Riverside-Orange County (Zook 2000, Malecki 2000, 2002, Townsend 2001). Table 4-5 shows a comparison of C/MSA Internet Ranking confirming that those cities that lead in terms of Internet backbone connectivity are the same C/MSAs that maintain top ranking for other types of Internet infrastructure and Internet activity (McIntee 2001). New York-Northern New Jersey-Long Island is nearly always included in the top five ranks when Internet infrastructure is measured (McIntee 2001, Kellerman 2002, Malecki & Gorman 2003). It appears that Atlanta has hedged New York-Northern New Jersey-Long Island for its position amongst the usual group, pushed back to the sixth rank Another surprise is the ranking of St.

Louis, 10. St. Louis is ranked 20th in terms of redundant connections. Table 4-5 shows that St. Louis is not included within the top 10 ranks of other types of Internet and telecommunication measurement either. Figure 4-8 illustrates the positive relationship between population and the top 25 nodes. Those nodes that top the ranking for the long-haul network also lead the ranking of cities in population. The map indicates those areas more densely populated are also the top ranked nodes for long-haul fiber. Kellerman (2002) suggests that the ranking of bandwidth measurement seems to approach the rank order of the population-based urban hierarchy.

Link Ranking

The results of powering the first matrix were also used to identify the top 12 links in the overall network (Table 4-6). The top 12 values within the resultant matrix were located, indicating corresponding cities that were linked by that particular value. A simple search was performed on the final product of each network multiplication scenario to identify the top 12 links within each network. After the links were identified, the values were converted to more manageable numbers. The method used is similar to the methods used in creating the node index. Each value was then divided by the lowest value of the top twelve links to create the index. The lowest valued link within the top 12 links of the network is between Dallas-Fort Worth and New York-Northern New Jersey-Long Island, with a value of 50016106024. This index is also an unweighted relative connectivity index (URCI). The value for the link connecting Dallas-Fort Worth and New York-Northern New Jersey-Long Island (50016106024) served as the starting point for the index. The value was reassigned as 1.0 to begin the index. Each link value was divided by the constant (50016106024); $URCI_i = \frac{X_i}{X_{min}}$, for i = 1...N, where N = 12. For example, to calculate the URCI value for the link between Chicago-Gary-Kenosha and Washington-Baltimore (link value 63901994600), $URCI_i = \frac{X_i}{X_{min}} = \frac{63901994600}{50016106024} = 1.278$. This value is 27% larger than X min.

Table 4-5. Comparison of C/MSA Internet rankings

FiberLit buildings (GeoTel Data 2003)	Colocation facilities (McIntee 2001)	Internet backbone (Townsend 2001)
New York-Northern New Jersey-Long Island	New York-Northern New Jersey-Long Island	Washington-Baltimore
Washington-Baltimore	Los Angeles-Riverside-Orange County	San Francisco-Oakland
San Francisco-Oakland	San Francisco-Oakland	Chicago-Gary-Kenosha
Los Angeles-Riverside-Orange County	Washington-Baltimore	New York-Northern New Jersey-Long Island
Dallas-Fort Worth	Dallas-Fort Worth	Dallas-Fort Worth
Chicago-Gary-Kenosha	Atlanta	Atlanta
Memphis	Boston-Worcester-Lawrence	Los Angeles-Riverside-Orange County
Houston-Galveston-Brazoria	Miami-Ft. Lauderdale	Seattle-Tacoma-Bremerton
Boston-Worcester-Lawrence	Chicago-Gary-Kenosha	Boston-Worcester-Lawrence
Atlanta	Seattle-Tacoma-Bremerton	Houston
Domain names (Zook 2000)	**Web development firms (Malecki 2000)**	**Cellular structure (Gorman & McIntee 2003)**
Los Angeles-Riverside-Orange County	San Francisco-Oakland	Los Angeles-Riverside-Orange County
New York-Northern New Jersey-Long Island	New York-Northern New Jersey-Long Island	New York-Northern New Jersey-Long Island
Washington-Baltimore	Los Angeles-Riverside-Orange County	San Francisco-Oakland
Chicago-Gary-Kenosha	Boston-Worcester-Lawrence	Dallas-Fort Worth
San Francisco-Oakland	Washington-Baltimore	Boston-Worcester-Lawrence
Boston-Worcester-Lawrence	Chicago-Gary-Kenosha	Philadelphia-Wilmington-Atlantic City
Orange County	Atlanta	Houston
San Jose	Seattle-Tacoma-Bremerton	Chicago-Gary-Kenosha
Atlanta	Miami-Ft. Lauderdale	Washington-Baltimore
Philadelphia-Wilmington-Atlantic City	San Diego	Denver

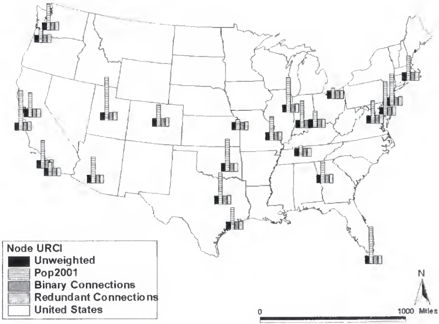


Figure 4-8. U.S. population (2000) density distribution and the top 25 nodes based on URCI values in the U.S. Internet backbone network.

It is interesting to note that the top three links in the original matrix each include Chicago-Gary-Kenosha as a node. Each of the cities displayed in Table 4-6 comprise the top six nodes in the network. Chicago-Gary-Kenosha ranks first in terms of Internet backbone connectivity, which helps to explain why the top three links in the overall network connect Chicago-Gary-Kenosha to other top cities. Townsend (2001) has suggested that Chicago-Gary-Kenosha is not considered in the top ranking of network cities, though it is a global city. In 2002, O'Kelly and Grubesic found Chicago to the most accessible city on the backbone network. This research suggests that Chicago-Gary-Kenosha is established as an Internet hub, rising amongst the network cities. In the past various C/MSAs outranked Chicago-Gary-Kenosha in terms of Internet and telecommunication activity, but based on this analysis, this city is outranked by none.

Table 4-6. Ranking of Internet backbone "links" in the U.S.

Rank	C/MSA	C/MSA	URIC
1	Chicago-Gary-Kenosha	Washington-Baltimore	1.278
2	Chicago-Gary-Kenosha	Dallas-Fort Worth	1.253
3	Chicago-Gary-Kenosha	San Francisco-Oakland-San Jose	1.173
4	Dallas-Fort Worth	Washington-Baltimore	1.160
5	Chicago-Gary-Kenosha	Atlanta	1.131
6	Chicago-Gary-Kenosha	New York-Northern New Jersey-Long Island	1.109
7	Washington-Baltimore	San Francisco-Oakland-San Jose	1.082
8	Dallas-Fort Worth	San Francisco-Oakland-San Jose	1.064
9	Washington-Baltimore	Atlanta	1.050
10	Dallas-Fort Worth	Atlanta	1.033
11	Washington-Baltimore	New York-Northern New Jersey-Long Island	1.023
12	Dallas-Fort Worth	New York-Northern New Jersey-Long Island	1.000

The single-node removal section of this chapter discusses the impact of removing Chicago-Gary-Kenosha from the network. As mentioned before, the diameter of the fully connected network is 11.The NCI is 99.499. The NCI will be used to calculate percentage change in the node removal scenarios. The removal scenarios will also be compared for change in rank, diameter, disconnects and NCI.

Table 4-7 displays basic measurements of the complete network. As mentioned before, the diameter of the fully connected network is 11 and the NCI is 99.499. The NCI will now be used to calculate percentage change in the node removal scenarios. The

removal scenarios will also be compared for change in rank, diameter, disconnects and NCI.

Table 4-7. Complete network measurements

Network connectivity	% Change	Diameter	Disconnects
99.499	N/A	11	Fully Connected

Basic statistical measures of the indexed ranking from the powering of the first matrix have been performed. Table 4-8 shows the results of these basic statistics. The descriptive statistical measures were performed on the 218 cities that create the long-haul fiber network. The URCI measures were used. The average URCI for the network is .456. The standard deviation shows how far the values are dispersed from that average, with a value of .630. Figure 4-9 illustrates the frequency of the URCI, indicating a positively skewed distribution. The cluster of scores at the lower end of the URCI scale are increasingly few. Roughly 50% (109) of the C/MSAs in the Internet backbone network have a URCI value higher than 1.0, indicating high connectivity to the rest of the network. The remaining half of the C/MSAs in the network have increasingly low URCI values, dispersed between .05 and .01.

Table 4-8. Descriptive statistics of U.S. Internet backbone network

Descriptive statistics of complete network	
Mean	0.456421
Median	0.244477
Standard deviation	0.629896
Minimum	0
Maximum	3.469602
Count	218

Node-Removal Scenarios

This section contains each removal scenario of the top twelve nodes that were identified in the powering of the original network. The results for each scenario include both a table of node rankings as well as a table of link rankings. However, due to space limitations, the majority of the results were consolidated into summary charts.

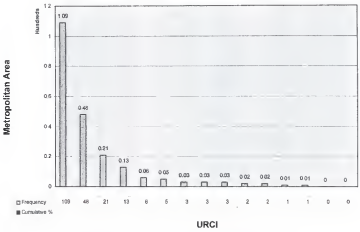


Figure 4-9. Connectivity distribution among 218 metropolitan areas

The individual removal of each of the top twelve nodes had little impact on the ranking shift of the top 5 nodes (Table 4-9). Six C/MSAs (Chicago-Gary-Kenosha, Atlanta, Washington-Baltimore, Dallas-Fort Worth, New York-Northern New Jersey-Long Island, San Francisco-Oakland-San Jose) remained consistently at the top of the network hierarchy of nodes for each of the dozen single-node removal scenarios that were performed.

Table 4-10 displays the NCI values, percentage change in connectivity, number of disconnects and diameter. Based on NCI values, the removal of Chicago-Gary-Kenosha has the greatest impact upon network connectivity, the value drops to 36.909. Chicago-Gary-Kenosha's removal causes the highest percentage change in connectivity at the 11th diameter, dropping 63%. The removal of Washington-Baltimore has the next most significant impact on the network, with a -58% change in connectivity. The network experienced a -683% change with the removal of Denver-Boulder-Greeley and 789% change with the removal of Sacramento-Yolo. Removing Denver-Boulder-Greeley and

Table 4-9. Ranking summary of 12 single-node removal scenarios for the U.S. Internet backbone network

	Node removal scenario												
	Original Network	Chicago	Washington-Baltimore	Dallas-Fort Worth	San Francisco	Atlanta	New York-Northern New Jersey	Kansas City	Denver	Los Angeles	St. Louis	Sacramento	Seattle
Chicago-Gary-Kenosha	1	X	1	1	1	1	2	1	1	1	1	1	1
Washington-Baltimore	2	2	X	2	3	3	3	2	2	2	2	3	2
Dallas-Fort Worth	3	1	2	X	4	4	1	3	4	3	3	2	3
San Francisco-Oakland-San Jose	4	3	3	4	X	2	5	6	5	5	6	5	5
Atlanta	5	4	4	5	2	X	4	5	3	4	4	4	4
New York-Northern New Jersey-Long Island	6	5	5	3	5	5	X	4	6	6	5	6	6

Sacramento-Yolo increased the network diameter to twelve. This means that the network was powered eleven times to achieve full connectivity. The rest of the network scenarios were powered ten times. This would cause the row totals to be much higher than if the network had only been multiplied ten times.

Table 4-10. Summary of 12 removal scenarios from the U.S. Internet backbone network

	Network Connectivity Index (NCI)	% Change	Diameter	Disconnects
Complete Network Values:	99.499	N/A	11	0
Network Removal Values:				
Chicago-Gary-Kenosha	36.909	-63.0	11	4
Washington-Baltimore	41.528	-58.3	11	0
Dallas-Fort Worth	44.506	-55.3	11	1
San Francisco-Oakland-San Jose	50.523	-49.0	11	3
Atlanta	50.677	-49.0	11	0
New York-Northern New Jersey-Long Island	52.979	-47.0	11	8
Kansas City	61.626	-38.0	11	0
Denver-Boulder-Greeley	68.284	-31.4	11	0
Los Angeles-Riverside-Orange County	66.954	-32.7	11	2
St. Louis	74.788	-24.8	11	0
Sacramento-Yolo	76.629	-22.9	11	0
Seattle-Tacoma-Bremerton	80.117	-19.4	11	1

The removal of New York-Northern New Jersey-Long Island causes the highest number of disconnects (Table 4-10). Eight cities are disconnected entirely from the network with the removal of the Big Apple. The removal of Chicago-Gary-Kenosha causes the second highest number of disconnects, four.

Chicago-Gary-Kenosha

When Chicago-Gary-Kenosha, which is the top ranking city in the complete network, is removed the top tier shifts little but the mid-ranking cities begin to show a more dramatic shift. With the removal of Chicago-Gary-Kenosha we see Tampa-St.Petersburg-Clearwater, Charlotte-Gastonia-Rock Hill, Los Angeles-Riverside-Orange County, and Orlando rise in ranking. In the original network they were not ranked in the

top 25. There are no radical shifts in ranking, which still holds Washington-Baltimore, Atlanta, San Francisco-Oakland-San Jose, New York-Northern New Jersey-Long Island, and Dallas-Fort Worth in the top five. The most critical link exists between Dallas-Fort Worth and Washington-Baltimore, which are also the top ranked nodes in the network. With the removal of Chicago-Gary-Kenosha, Dallas-Fort Worth rises to the top of the hierarchy. The top three links in this scenario connect Dallas-Fort Worth to other top ranking cities. The NCI with Chicago-Gary-Kenosha removed, 36.909, is dramatically lower than the original NCI value. Of the 12 nodes removed from the network, this is the highest drop in NCI value. As mentioned earlier in the text, the connectivity of the network without Chicago-Gary-Kenosha changes by -63%. This large drop in connectivity could be due to the importance of this node's central location in the U.S. Internet backbone network. This is consistent with previous Internet backbone research claiming that the central location makes it one of the most accessible cities (O'Kelly & Grubesic 2002) . Reviewing Table 4-5, it is apparent that Chicago-Gary-Kenosha has consistently ranked within the top 10 of various types of Internet infrastructure and activity.

Washington-Baltimore

When Washington-Baltimore, which ranks second in the overall network, is removed from the network, no cities are disconnected. The absence of disconnects is probably due to the close proximity of this node to many other nodes, allowing for more direct connections to other C/MSAs (such as Boston, New York, Philadelphia, Atlanta). Table 4-4 confirms that Washington-Baltimore does have more redundant than another other C/MSA (245 links) . In 1999 Washington-Baltimore ranked first in both number of links and total bandwidth, but the dominance of links is higher than in bandwidth (Kellerman 2002, p. 145). The network ranking shifts very little. However, the network connectivity drops significantly to 41.522, a -58% change from the original network. The

Table 4-11. U.S. city ranking for the Internet backbone network based on the removal of Chicago-Gary-Kenosha

Rank	CMSA	URCI
1	Dallas-Fort Worth	1.290
2	Washington-Baltimore	1.272
3	Atlanta	1.155
4	San Francisco-Oakland-San Jose	1.149
5	New York-Northern New Jersey-Long Island	1 075
6	Los Angeles-Riverside-Orange County	1.020
7	Kansas City	0.920
8	Denver-Boulder-Greeley	0.861
9	Sacramento-Yolo	0.831
10	Houston-Galveston-Brazoria	0.755
11	St. Louis	0.651
12	Phoenix-Mesa	0.604
13	San Diego	0.599
14	Seattle-Tacoma-Bremerton	0.565
15	Miami-Fort Lauderdale	0.563
16	Boston-Worcester-Lawrence	0.556
17	Pennsauken	0.500
18	Cleveland-Akron	0.479
19	Nashville	0.476
20	Portland-Salem	0.435
21	Las Vegas	0.434
22	Salt Lake City-Ogden	0.428
23	Orlando	0.426
24	Charlotte-Gastonia-Rock Hill	0.418
25	Tampa-St.Petersburg-Clearwater	0 418

Note: NCI 36.909; -63.0% change; diameter 11; disconnects: Oakbrook, Rolling Meadows, Southfield, Glenview

Table 4-12 Ranking of links based on the removal of Chicago from the U.S. Internet backbone network

Rank	C/MSA	C/MSA	URIC
1	Dallas-Fort Worth	Washington-Baltimore	0.507
2	Dallas-Fort Worth	Atlanta	0.461
3	Dallas-Fort Worth	San Francisco-Oakland-San Jose	0.457
4	Washington-Baltimore	Atlanta	0 454
5	Washington-Baltimore	San Francisco-Oakland-San Jose	0.449
6	Dallas-Fort Worth	New York-Northern New Jersey-Long Island	0.423
7	New York-Northern New Jersey-Long Island	Washington-Baltimore	0.421
8	Dallas-Fort Worth	Los Angeles-Riverside-Orange County	0.405
9	Atlanta	San Francisco-Oakland-San Jose	0.399
10	Los Angeles-Riverside-Orange County	Washington-Baltimore	0.397
11	San Francisco-Oakland-San Jose	New York-Northern New Jersey-Long Island	0.383
12	Atlanta	New York-Northern New Jersey-Long Island	0.375

diameter remains at 11. Washington, DC, serves as more than the center of government but also a center of high-tech activity (Kellerman 2002, p.145). Washington-Baltimore has more high-tech employees than Federal employees (Grubesic & O'Kelly 2002, Moss & Townsend 2000). The most critical link to the overall network exists between Chicago-Gary-Kenosha and Dallas-Fort Worth, which are similar to the first node-removal scenario and are the top two nodes in the ranking with Washington-Baltimore removed. The central location of both Chicago-Gary-Kenosha and Dallas-Fort Worth cannot be overlooked. It appears that geographic centrality is important to the connectivity of the network, particularly with the removal of Washington-Baltimore. This implies that though Washington-Baltimore is not located centrally, it serves as a central node for the network With it's absence, nodes that are centrally located geographically rise in rank. becoming more important to the network's connectivity.

Dallas-Fort Worth

When Dallas-Fort Worth is disconnected from the network, Chicago-Gary-Kenosha remains at the top of the network with Washington-Baltimore ranking second (Table 4-15). Only one city, Amarillo, is disconnected from the network in this scenario. With the loss of Dallas-Fort Worth, the NCI also drops dramatically, to 44.506. This is a -55% change in connectivity compared to the complete network. Chicago-Gary-Kenosha and Washington-Baltimore share the most critical link of the network (Table 4-16). Kellerman (2002) has suggested that the central geographic location of both Dallas-Fort Worth and Atlanta explains their local interconnectivity with other cities. Dallas-Fort Worth ranks fourth in terms of redundant connections. Dallas-Fort Worth was also home to one of the original interconnection points of the Internet, MAE-Dallas. Metropolitan Area Exchanges (MAE) are owned and operated by MCI-WorldCom (Malecki & Gorman 2001, eds Leinbach & Brunn). A MAE, like other types of interconnection points, physically connects Internet backbones within its facility. Interconnection is crucial to the

Table 4-13. U.S. city ranking for the Internet backbone network based on the removal of Washington-Baltimore

Rank	CMSA	URCI
1	Chicago-Gary-Kenosha	1.527
2	Dallas-Fort Worth	1.393
3	San Francisco-Oakland-San Jose	1.319
4	Atlanta	1.179
5	New York-Northern New Jersey-Long Island	1.179
6	Denver-Boulder-Greeley	1.112
7	Kansas City	1.061
8	Los Angeles-Riverside-Orange County	0.942
9	Seattle-Tacoma-Bremerton	0.810
10	Sacramento-Yolo	0.806
11	St. Louis	0.788
12	Houston-Galveston-Brazoria	0.645
13	San Diego	0.624
14	Phoenix-Mesa	0.621
15	Indianapolis	0.615
16	Boston-Worcester-Lawrence	0.593
17	Cleveland-Akron	0.591
18	Miami-Fort Lauderdale	0.580
19	Tulsa	0.577
20	Nashville	0.520
21	Salt Lake City-Ogden	0.512
22	Portland-Salem	0.489
23	Las Vegas	0.475
24	Pennsauken	0.442
25	Orlando	0.437

Note: NCI 41.522; -58.0 % change, diameter 11; no disconnects

Table 4-14. Ranking of links based on the removal of Washington-Baltimore from the U.S. Internet backbone network

Rank	C/MSA	C/MSA	URCI
1	Chicago-Gary-Kenosha	Dallas-Fort Worth	0.584
2	Chicago-Gary-Kenosha	San Francisco-Oakland-San Jose	0.555
3	Dallas-Fort Worth	San Francisco-Oakland-San Jose	0.503
4	Chicago-Gary-Kenosha	New York-Northern New Jersey-Long Island	0.496
5	Chicago-Gary-Kenosha	Atlanta	0.490
6	Chicago-Gary-Kenosha	Denver-Boulder-Greeley	0.465
7	Dallas-Fort Worth	Atlanta	0.451
8	Dallas-Fort Worth	New York-Northern New Jersey-Long Island	0.444
9	Chicago-Gary-Kenosha	Kansas City	0.443
10	San Francisco-Oakland-San Jose	New York-Northern New Jersey-Long Island	0.428
11	Dallas-Fort Worth	Denver-Boulder-Greeley	0.423
12	San Francisco-Oakland-San Jose	Atlanta	0.416

Table 4-15. U.S. city ranking for the Internet backbone network based on the removal of Dallas-Fort Worth

Rank	CMSA	URCI
1	Chicago-Gary-Kenosha	1.663
2	Washington-Baltimore	1.505
3	New York-Northern New Jersey-Long Island	1.489
4	San Francisco-Oakland-San Jose	1.367
5	Atlanta	1.276
6	Kansas City	1.066
7	Denver-Boulder-Greeley	0.974
8	Los Angeles-Riverside-Orange County	0.960
9	Sacramento-Yolo	0.950
10	Cleveland-Akron	0.882
11	Seattle-Tacoma-Bremerton	0.861
12	Boston-Worcester-Lawrence	0.856
13	St. Louis	0.787
14	Indianapolis	0.710
15	Miami-Fort Lauderdale	0.618
16	Portland-Salem	0.562
17	Houston-Galveston-Brazoria	0.556
18	Cincinnati-Hamilton	0.537
19	Salt Lake City-Ogden	0.523
20	Pennsauken	0.507
21	San Diego	0.476
22	Charlotte-Gastonia-Rock Hill	0.474
23	Louisville	0.466
24	Raliegh-Durham-Chapel Hill	0.462
25	Tulsa	0.451

Note: NCI 44.506; -55.0 % change; diameter 11; disconnects Amarillo

exchange of data between networks. The central location of Dallas-Fort Worth and the necessity of interconnection explain the importance of this node to the overall network.

San Francisco-Oakland-San Jose

Honolulu is completely disconnected from the nation's long-haul fiber network with the removal of San Francisco-Oakland-San Jose. Bellevue & Redwood city are also disconnected. Like the other top cities in the connectivity hierarchy, the removal of San Francisco-Oakland-San Jose causes the network to experience a heavy drop in connectivity, to 50. 523. This is a 49% decrease in connectivity. There is little shift in the

network's top ranks. Chicago-Gary-Kenosha, Atlanta, Washington-Baltimore, Dallas and New York-Northern New Jersey-Long Island fill the top five ranks (Table 4-17). Table 4.18 shows little surprise in terms of the most prominent links in the network. As before, the link between Chicago-Gary-Kenosha and Washington-Baltimore remains the most critical link to the network (Table 4-18). San Francisco-Oakland-San Jose has long been a key connection point for data transfer and hub of Internet activity Several major interconnection points are located in San Francisco-Oakland-San Jose. In 1994, one of the original Network Access Points (NAP) was built by the National Science Foundation in San Francisco. At that time, San Francisco had experienced a large growth in Internet infrastructure, thus prompting the demand for interconnection. The NAPs were created to facilitate interconnection. The other NAPs were established in Chicago-Gary-Kenosha and New York (Malecki & Gorman 2001). San Francisco-Oakland-San Jose was a key player in the national data transfer network decades before the commercial Internet emerged (Figure 4-10, courtesy of Townsend, personal communication, 2003) As the Internet continues to grow, San Francisco-Oakland-San Jose maintains its dominance as a critical node (Table 4-5).

Atlanta

In each of the removal scenarios of the top twelve nodes, Atlanta consistently ranked within the top five most crucial nodes in the networks (Table 4-9). Atlanta's central location (in terms of its position in the network, not physical location), like Dallas-Fort Worth, has been crucial in boosting this city to a place of importance in the ranking of the Internet backbone network. Unlike Dallas-Fort Worth, Atlanta became a member of NSFNet later in the development of the network (Figure 4-11). Today however, Atlanta ranks sixth in terms of binary links. The NCI dropped to 50.677 (-49%) when Atlanta was removed, confirming the importance of this centrally located node.

Table 4-16. Ranking of links based on the removal of Dallas-Fort Worth from the U.S. Internet backbone network

Rank	C/MSA	C/MSA	URCI
1	Chicago-Gary-Kenosha	Washington-Baltimore	0.643
2	Chicago-Gary-Kenosha	New York-Northern New Jersey-Long Island	0.638
3	Chicago-Gary-Kenosha	San Francisco-Oakland-San Jose	0.584
4	Washington-Baltimore	New York-Northern New Jersey-Long Island	0.573
5	Chicago-Gary-Kenosha	Atlanta	0.539
6	Washington-Baltimore	San Francisco-Oakland-San Jose	0.524
7	New York-Northern New Jersey-Long Island	San Francisco-Oakland-San Jose	0.523
8	Washington-Baltimore	Atlanta	0.490
9	New York-Northern New Jersey-Long Island	Atlanta	0.479
10	Chicago-Gary-Kenosha	Kansas City	0.452
11	San Francisco-Oakland-San Jose	Atlanta	0.434
12	Chicago-Gary-Kenosha	Denver-Boulder-Greeley	0.412

Table 4-17. U.S. city ranking for the Internet backbone network based on the removal of San Francisco-Oakland-San Jose

Rank	CMSA	URCI
1	Chicago-Gary-Kenosha	1.788
2	Atlanta	1.699
3	Washington-Baltimore	1.694
4	Dallas-Fort Worth	1.633
5	New York-Northern New Jersey-Long Island	1.388
6	Kansas City	1.195
7	Los Angeles-Riverside-Orange County	1.042
8	Denver-Boulder-Greeley	1.009
9	Miami-Fort Lauderdale	0.917
10	Houston-Galveston-Brazoria	0.914
11	St. Louis	0.911
12	Cleveland-Akron	0.908
13	Boston-Worcester-Lawrence	0.908
14	Sacramento-Yolo	0.761
15	Seattle-Tacoma-Bremerton	0.694
16	Nashville	0.690
17	Pennsauken	0.664
18	Phoenix-Mesa	0.635
19	Indianapolis	0.608
20	Cincinnati-Hamilton	0.599
21	Tampa-St.Petersburg-Clearwater	0.592
22	Orlando	0.587
23	Charlotte-Gastonia-Rock Hill	0.584
24	Raliegh-Durham-Chapel Hill	0.571
25	San Diego	0.539

Note: NCI 50.523; -49.0 % change; diameter 11; disconnects Bellevue, Redwood City, Honolulu

Table 4-18. Ranking of links based on the removal of San Francisco-Oakland-San Jose from the U.S. Internet backbone network

Rank	C/MSA	C/MSA	URCI
1	Washington-Baltimore	Chicago-Gary-Kenosha	0.684
2	Chicago-Gary-Kenosha	Atlanta	0.683
3	Chicago-Gary-Kenosha	Dallas-Fort Worth	0.657
4	Atlanta	Washington-Baltimore	0.649
5	Atlanta	Dallas-Fort Worth	0.627
6	Washington-Baltimore	Dallas-Fort Worth	0.622
7	Chicago-Gary-Kenosha	New York-Northern New Jersey-Long Island	0.560
8	Washington-Baltimore	New York-Northern New Jersey-Long Island	0.527
9	Atlanta	New York-Northern New Jersey-Long Island	0.524
10	Dallas-Fort Worth	New York-Northern New Jersey-Long Island	0.503
11	Chicago-Gary-Kenosha	Kansas City	0.478
12	Atlanta	Kansas City	0.452

New York-Northern New Jersey-Long Island

The removal of the Big Apple has less impact on the network than Chicago-Gary-Kenosha, Washington-Baltimore, Dallas-Fort Worth, San Francisco-Oakland-San Jose, or Atlanta. Considering that New York-Northern New Jersey-Long Island ranks second in binary links this is a bit surprising. New York-Northern New Jersey-Long Island ranks first in terms of fiber-lit buildings and colocation facilities (Table 4-6). This node ranks second in measurement of domain names and web development firms (Table 4-6). Eight cities are disconnected with the loss of New York-Northern New Jersey-Long Island (Table 4-21). This is more disconnects than any other node-removal Table 4-10. These results are in direct contrast to Malecki and Gorman's contention that New York is not included in the exclusive group of metropolitan areas that are most important to the Internet backbone network. Malecki and Gorman (2001, eds. Brunn & Leinbach) suggest that New York-Northern New Jersey-Long Island is not in the top of a four-tier system based on binary connectivity for 2000. Washington-Baltimore, Chicago-Gary-Kenosha, San Francisco-Oakland-San Jose comprise the top tier, with New York-Northern New Jersey-Long Island, Atlanta and Dallas-Fort Worth comprising the second tier In this removal scenario, the NCI has dropped to 52.979, showing a 47% drop in connectivity (Table 4-10). Dallas-Fort Worth leads the ranking with the removal of New York-Northern

Figure 4-10. Cities connected to Advanced Research Project Network (ARPANET) in 1971 (Courtesy of Townsend 2003)

Table 4-19. U S. city ranking based on the removal of Atlanta from the Internet backbone network

Rank	CMSA	URCI
1	Chicago-Gary-Kenosha	1 887
2	San Francisco-Oakland-San Jose	1.783
3	Washington-Baltimore	1.654
4	Dallas-Fort Worth	1.640
5	New York-Northern New Jersey-Long Island	1.546
6	Denver-Boulder-Greeley	1.367
7	Kansas City	1.327
8	Los Angeles-Riverside-Orange County	1.177
9	Sacramento-Yolo	1.156
10	St. Louis	1.080
11	Seattle-Tacoma-Bremerton	1.018
12	Cleveland-Akron	0.913
13	Indianapolis	0.803
14	Boston-Worcester-Lawrence	0.798
15	San Diego	0.766
16	Phoenix-Mesa	0.744
17	Houston-Galveston-Brazoria	0.714
18	Tulsa	0.713
19	Pennsauken	0.692
20	Salt Lake City-Ogden	0.651
21	Portland-Salem	0.636
22	Miami-Fort Lauderdale	0.588
23	Las Vegas	0.586
24	Cincinnati-Hamilton	0.584
25	Stockton-Lodi	0.540

Note: NCI 50.677; -49.0 % change; diameter 11; no disconnects

Table 4-20. Ranking of links based on the removal of Atlanta from the U.S. Internet backbone network

Rank	C/MSA	C/MSA	URCI
1	Chicago-Gary-Kenosha	San Francisco-Oakland-San Jose	0.759
2	Chicago-Gary-Kenosha	Washington-Baltimore	0.703
3	Chicago-Gary-Kenosha	Dallas-Fort Worth	0.695
4	San Francisco-Oakland-San Jose	Washington-Baltimore	0.662
5	Dallas-Fort Worth	San Francisco-Oakland-San Jose	0.658
6	Chicago-Gary-Kenosha	New York-Northern New Jersey-Long Island	0.657
7	San Francisco-Oakland-San Jose	New York-Northern New Jersey-Long Island	0.618
8	Washington-Baltimore	Dallas-Fort Worth	0.607
9	Chicago-Gary-Kenosha	Denver-Boulder-Greeley	0.576
10	New York-Northern New Jersey-Long Island	Washington-Baltimore	0.572
11	Dallas-Fort Worth	New York-Northern New Jersey-Long Island	0.563
12	Chicago-Gary-Kenosha	Kansas City	0.560

Figure 4-11. Cities connected to the Advanced Research Project Network (ARPANET) in 1980 (Courtsey of Townsend 2003)

New Jersey-Long Island, rather than the usual leader, Chicago-Gary-Kenosha. The highest ranked link also shifts. The link between Chicago-Gary-Kenosha and Washington-Baltimore has been first rank in the majority of the other node removal scenarios. However, with this scenario the most prominent link is the connection between Dallas-Fort Worth and Chicago-Gary-Kenosha (Table 4-22). The second most prominent link is between Dallas-Fort Worth and Washington-Baltimore. The third link, is between Chicago-Gary-Kenosha and Washington-Baltimore, which usually ranks first (Table 4-22).

Kansas City

Surprisingly, Kansas City is ranked seventh amongst the nodes in the Internet backbone network. Ranking eleventh in terms of redundant connections, the central location of this node might increase its importance to the overall network. Based on data from 1999, Kansas City had twice the number of binary connections expected for it's population (Kellerman 2002). Network connectivity with the loss of Kansas City drops to 61.626, a 38% change in connectivity. No disconnects occur with the loss of this particular node. The ranking listed in Tables 4-23 and 4-24 shows little change in the network hierarchy of nodes and links.

Denver-Boulder-Greeley

Denver-Boulder-Greeley was the eighth ranked node in the network. When Denver-Boulder-Greeley was removed from the network the NCI drops to 68.284, a 31.4 % decrease in network connectivity. No disconnects occurred, and the ranking looks predictable. The URCI values for each of the top five nodes are close (Table 4-25). Table 4-26 shows the link ranking for the network, with the prominent link between Chicago-Gary-Kenosha and Washington-Baltimore ranking first.

Table 4-21. U.S. city ranking based on the removal of New York-Northern New Jersey-Long Island from the Internet backbone network

Rank	CMSA	URCI
1	Dallas-Fort Worth	1 946
2	Chicago-Gary-Kenosha	1.917
3	Washington-Baltimore	1.757
4	Atlanta	1.648
5	San Francisco-Oakland-San Jose	1.593
6	Kansas City	1.460
7	Los Angeles-Riverside-Orange County	1.216
8	Denver-Boulder-Greeley	1.205
9	St. Louis	1.150
10	Houston-Galveston-Brazoria	1.001
11	Miami-Fort Lauderdale	0.949
12	Sacramento-Yolo	0.937
13	Seattle-Tacoma-Bremerton	0.807
14	San Diego	0.770
15	Phoenix-Mesa	0.767
16	Tulsa	0.745
17	Nashville	0.736
18	Cleveland-Akron	0.729
19	Boston-Worcester-Lawrence	0.719
20	Indianapolis	0.642
21	Orlando	0.607
22	Cincinnati-Hamilton	0 606
23	Tampa-St.Petersburg-Clearwater	0.594
24	Pennsauken	0.591
25	Charlotte-Gastonia-Rock Hill	0.584

Note: NCI 52.979; -47.0 % change ;diameter 11; disconnects 8 cities: Garden City, Cedar Knolls, Bohemia, Hackensack, Freehold, Wayne, Bridgeport, Rochelle Park

Table 4-22 Ranking of links based on the removal of New York-Northern New Jersey-Long Island from the U.S. Internet backbone network

Rank	C/MSA	C/MSA	URCI
1	Dallas-Fort Worth	Chicago-Gary-Kenosha	0.805
2	Dallas-Fort Worth	Washington-Baltimore	0.736
3	Chicago-Gary-Kenosha	Washington-Baltimore	0.725
4	Dallas-Fort Worth	Atlanta	0.689
5	Atlanta	Chicago-Gary-Kenosha	0.676
6	Dallas-Fort Worth	San Francisco-Oakland-San Jose	0.667
7	Chicago-Gary-Kenosha	San Francisco-Oakland-San Jose	0.657
8	Washington-Baltimore	Atlanta	0 621
9	Dallas-Fort Worth	Kansas City	0.609
10	Chicago-Gary-Kenosha	Kansas City	0.600
11	Washington-Baltimore	San Francisco-Oakland-San Jose	0.597
12	Atlanta	San Francisco-Oakland-San Jose	0.553

Table 4-23. U.S. city ranking based on the removal of Kansas City from the Internet backbone network

Rank	CMSA	URCI
1	Chicago-Gary-Kenosha	2.134
2	Washington-Baltimore	2.006
3	Dallas-Fort Worth	1.909
4	New York-Northern New Jersey-Long Island	1.851
5	Atlanta	1.809
6	San Francisco-Oakland-San Jose	1.778
7	Los Angeles-Riverside-Orange County	1.346
8	Denver-Boulder-Greeley	1.292
9	Sacramento-Yolo	1.221
10	Cleveland-Akron	1.086
11	St. Louis	1.085
12	Boston-Worcester-Lawrence	1.061
13	Houston-Galveston-Brazoria	1.057
14	Miami-Fort Lauderdale	1.013
15	Seattle-Tacoma-Bremerton	0.931
16	Indianapolis	0.886
17	San Diego	0.831
18	Phoenix-Mesa	0.817
19	Nashville	0.764
20	Cincinnati-Hamilton	0.707
21	Pennsauken	0.678
22	Portland-Salem	0.676
23	Charlotte-Gastonia-Rock Hill	0.642
24	Salt Lake City-Ogden	0.638
25	Tulsa	0.635

Note: NCI: 61.626; -38.0 % change; diameter 11; no disconnects

Table 4-24. Ranking of links based on the removal of Kansas City from the U.S. Internet backbone network

Rank	C/MSA	C/MSA	URCI
1	Chicago-Gary-Kenosha	Washington-Baltimore	0.791
2	Chicago-Gary-Kenosha	Dallas-Fort Worth	0.750
3	Chicago-Gary-Kenosha	New York-Northern New Jersey-Long Island	0.731
4	Chicago-Gary-Kenosha	Atlanta	0.708
5	Washington-Baltimore	Dallas-Fort Worth	0.704
6	Chicago-Gary-Kenosha	San Francisco-Oakland-San Jose	0.700
7	Washington-Baltimore	New York-Northern New Jersey-Long Island	0.683
8	Washington-Baltimore	Atlanta	0.668
9	Washington-Baltimore	San Francisco-Oakland-San Jose	0.654
10	Dallas-Fort Worth	New York-Northern New Jersey-Long Island	0.645
11	Atlanta	Dallas-Fort Worth	0.636
12	San Francisco-Oakland-San Jose	Dallas-Fort Worth	0.623

Table 4-23. U.S. city ranking based on the removal of Kansas City from the Internet backbone network

Rank	CMSA	URCI
1	Chicago-Gary-Kenosha	2.134
2	Washington-Baltimore	2.006
3	Dallas-Fort Worth	1.909
4	New York-Northern New Jersey-Long Island	1.851
5	Atlanta	1.809
6	San Francisco-Oakland-San Jose	1.778
7	Los Angeles-Riverside-Orange County	1.346
8	Denver-Boulder-Greeley	1.292
9	Sacramento-Yolo	1.221
10	Cleveland-Akron	1.086
11	St. Louis	1.085
12	Boston-Worcester-Lawrence	1 061
13	Houston-Galveston-Brazoria	1.057
14	Miami-Fort Lauderdale	1.013
15	Seattle-Tacoma-Bremerton	0.931
16	Indianapolis	0.886
17	San Diego	0.831
18	Phoenix-Mesa	0.817
19	Nashville	0.764
20	Cincinnati-Hamilton	0.707
21	Pennsauken	0.678
22	Portland-Salem	0.676
23	Charlotte-Gastonia-Rock Hill	0.642
24	Salt Lake City-Ogden	0.638
25	Tulsa	0.635

Note: NCI: 61.626; -38.0 % change; diameter 11; no disconnects

Table 4-24. Ranking of links based on the removal of Kansas City from the U S. Internet backbone network

Rank	C/MSA	C/MSA	URCI
1	Chicago-Gary-Kenosha	Washington-Baltimore	0.791
2	Chicago-Gary-Kenosha	Dallas-Fort Worth	0.750
3	Chicago-Gary-Kenosha	New York-Northern New Jersey-Long Island	0.731
4	Chicago-Gary-Kenosha	Atlanta	0.708
5	Washington-Baltimore	Dallas-Fort Worth	0.704
6	Chicago-Gary-Kenosha	San Francisco-Oakland-San Jose	0.700
7	Washington-Baltimore	New York-Northern New Jersey-Long Island	0.683
8	Washington-Baltimore	Atlanta	0 668
9	Washington-Baltimore	San Francisco-Oakland-San Jose	0.654
10	Dallas-Fort Worth	New York-Northern New Jersey-Long Island	0.645
11	Atlanta	Dallas-Fort Worth	0.636
12	San Francisco-Oakland-San Jose	Dallas-Fort Worth	0 623

Table 4-25. U.S. city ranking based on the removal of Denver-Boulder-Greeley from the Internet backbone network

Rank	CMSA	URCI
1	Chicago-Gary-Kenosha	2.463
2	Washington-Baltimore	2.430
3	Atlanta	2.212
4	Dallas-Fort Worth	2.189
5	San Francisco-Oakland-San Jose	1.959
6	New York-Northern New Jersey-Long Island	1.938
7	Kansas City	1.632
8	Los Angeles-Riverside-Orange County	1.577
9	St. Louis	1.393
10	Cleveland-Akron	1 204
11	Boston-Worcester-Lawrence	1.203
12	Miami-Fort Lauderdale	1.163
13	Houston-Galveston-Brazoria	1.158
14	Sacramento-Yolo	1.157
15	Seattle-Tacoma-Bremerton	0.985
16	Indianapolis	0.979
17	Pennsauken	0.888
18	Nashville	0.885
19	Tulsa	0.856
20	Cincinnati-Hamilton	0.797
21	San Diego	0.756
22	Charlotte-Gastonia-Rock Hill	0.746
23	Raliegh-Durham-Chapel Hill	0.725
24	Orlando	0.718
25	Tampa-St.Petersburg-Clearwater	0.716

Note: NCI 68 284; -31.4 % change; diameter 11; no disconnects

Los Angeles-Riverside-Orange County

Los Angeles-Riverside-Orange County ranks ninth in terms of connectivity for the Long-haul fiber optic network. This is surprisingly low compared to Los Angeles-Riverside-Orange County's placement in other measures of Internet rankings. Los Angeles-Riverside-Orange County ranks fourth in Fiber-lit buildings, second in colocation facilities, first in domain names, third in web development firms, and first in cellular structures (see Table 4-5). Los Angeles-Riverside-Orange County ranks seventh in redundant links and seventh in binary links (see table 4-2 & 4-4). It appears that Los Angeles-Riverside-Orange County is weaker in terms of long-haul binary links than other types of infrastructures and activity involving data transfer. Kellerman (2002) ranks Los

Table 4-26. Ranking of links based on the removal of Denver-Boulder-Greeley from the U.S Internet backbone network

Rank	C/MSA	C/MSA	URCI
1	Chicago-Gary-Kenosha	Washington-Baltimore	0.944
2	Chicago-Gary-Kenosha	Atlanta	0.856
3	Chicago-Gary-Kenosha	Dallas-Fort Worth	0.849
4	Washington-Baltimore	Atlanta	0.847
5	Washington-Baltimore	Dallas-Fort Worth	0.838
6	Atlanta	Dallas-Fort Worth	0.766
7	Chicago-Gary-Kenosha	San Francisco-Oakland-San Jose	0.762
8	Chicago-Gary-Kenosha	New York-Northern New Jersey-Long Island	0.757
9	Washington-Baltimore	San Francisco-Oakland-San Jose	0.751
10	New York-Northern New Jersey-Long Island	Washington-Baltimore	0.744
11	San Francisco-Oakland-San Jose	Atlanta	0.678
12	Dallas-Fort Worth	San Francisco-Oakland-San Jose	0.676

Angeles-Riverside-Orange County ninth in terms of intercontinental backbones, which may explain why Los Angeles-Riverside-Orange County maintains high ranking in Internet measurements (Table 4.5) though she may lag in terms of binary connections (Table 4-2). Two cities are disconnected from the network with the removal of Los Angeles-Riverside-Orange County, and network connectivity is reduced by 33%. The NCI is 66.954, with the usual group comprising the top of the ranking (Table 4-27). As the usual group comprises the top of the node ranking, the same holds true for the ranking of links (Table 4-28). The connection between Chicago and the other five nodes comprising the top of the nodal ranking Washington-Baltimore, Dallas-Forth Worth, Atlanta, San Francisco-Oakland-San Jose & New York-Northern New Jersey-Long Island create the top five links (Table 4-28).

St. Louis

The tenth ranked node in the network is St. Louis. St. Louis is not ranked among the top 15 in terms of binary links (Table 4-2). St. Louis does not rank high in terms of other types of Internet infrastructure or activity either (Table 4-5).

Table 4-27. U.S. city ranking based on the removal of Los Angeles-Riverside-Orange County from the Internet backbone network

Rank	CMSA	URCI
1	Chicago-Gary-Kenosha	2.526
2	Washington-Baltimore	2.170
3	Dallas-Fort Worth	2.094
4	Atlanta	1.943
5	San Francisco-Oakland-San Jose	1.918
6	New York-Northern New Jersey-Long Island	1.883
7	Kansas City	1.633
8	Denver-Boulder-Greeley	1.534
9	St. Louis	1.367
10	Cleveland-Akron	1.181
11	Boston-Worcester-Lawrence	1.164
12	Seattle-Tacoma-Bremerton	1.163
13	Sacramento-Yolo	1.146
14	Miami-Fort Lauderdale	1.099
15	Indianapolis	0.984
16	Houston-Galveston-Brazoria	0.961
17	Pennsauken	0.865
18	Tulsa	0.863
19	Nashville	0.843
20	Cincinnati-Hamilton	0.785
21	San Diego	0.731
22	Portland-Salem	0.712
23	Charlotte-Gastonia-Rock Hill	0.688
24	Louisville	0.687
25	Phoenix-Mesa	0.686

Note: NCI 66.954; -33.0 % change; diameter 11; disconnects: Sherman Oaks, Gardena

Table 4-28. Ranking of links based on the removal of Los Angeles-Riverside-Orange County from the U.S. Internet backbone network

Rank	C/MSA	C/MSA	URCI
1	Chicago-Gary-Kenosha	Washington-Baltimore	0.935
2	Chicago-Gary-Kenosha	Dallas-Fort Worth	0.901
3	Chicago-Gary-Kenosha	Atlanta	0.831
4	Chicago-Gary-Kenosha	San Francisco-Oakland-San Jose	0.826
5	Chicago-Gary-Kenosha	New York-Northern New Jersey-Long Island	0.811
6	Washington-Baltimore	Dallas-Fort Worth	0.769
7	Washington-Baltimore	Atlanta	0.714
8	Washington-Baltimore	San Francisco-Oakland-San Jose	0.702
9	Chicago-Gary-Kenosha	Kansas City	0.699
10	Washington-Baltimore	New York-Northern New Jersey-Long Island	0.691
11	Dallas-Fort Worth	Atlanta	0.690
12	Dallas-Fort Worth	San Francisco-Oakland-San Jose	0.678

Currently, St. Louis ranks 20th in terms of redundant links (Table 4-4). When St. Louis is removed from the network, no cities are disconnected. Network connectivity drops to 74.788, only 25 % change. Little shift occurs in the network, with node or link ranking (Table 4-29, Table 4-30).

Sacramento-Yolo

Sacramento-Yolo ranks eleventh in the network. Very little has been published on the infrastructure, or lack there of in Sacramento-Yolo. One explanation for the placement of Sacramento in this ranking is its adjacency to the San Francisco Bay area (a hub of Internet infrastructure and activity). Sacramento-Yolo ranks 18^{th} in terms of redundant links (Table 4-4). The NCI is 76 629, a -22.9 % drop in connectivity. Only one city was disconnected from the network. Table 4-31 and 4-32 show little shift in link or node rankings.

Seattle-Tacoma-Bremerton

Rounding off the node removal scenarios is Seattle-Tacoma-Bremerton. Seattle-Tacoma-Bremerton is ranked 15th in terms of binary links (see Table 4-2) and 13^{th} in terms of redundant links (see Table 4-4). By 1991, the Department of Defense had disabled the ARPANET and it was replaced with the National Science Foundation Network (NSFNET). Seattle-Tacoma-Bremerton was connected to the NSFNet backbone during this time-period. Figure 4-12 illustrates the NSF backbone in 1991. The NCI change with the removal of Seattle-Tacoma-Bremerton is only 20%. Anchorage is completely disconnected from the long haul fiber optic network with the removal of Seattle-Tacoma-Bremerton and the network ranking is standard (Table 4-33, Table 4-34).

Pair Removals

Pair removal scenarios were performed on each of the possible 66 pairings of the top twelve nodes by ranked URCI. A total of 72 pair-removal scenarios were performed,

Table 4-29. U.S. city ranking based on the removal of St. Louis from the Internet backbone network

Rank	CMSA	URCI
1	Chicago-Gary-Kenosha	2 557
2	Washington-Baltimore	2.414
3	Dallas-Fort Worth	2.324
4	Atlanta	2.247
5	New York-Northern New Jersey-Long Island	2.217
6	San Francisco-Oakland-San Jose	2.188
7	Kansas City	1.783
8	Los Angeles-Riverside-Orange County	1.766
9	Denver-Boulder-Greeley	1.760
10	Sacramento-Yolo	1.499
11	Seattle-Tacoma-Bremerton	1.296
12	Houston-Galveston-Brazoria	1.269
13	Boston-Worcester-Lawrence	1.259
14	Cleveland-Akron	1.227
15	Miami-Fort Lauderdale	1.195
16	San Diego	1.022
17	Phoenix-Mesa	1 009
18	Pennsauken	0.954
19	Indianapolis	0.903
20	Portland-Salem	0.820
21	Salt Lake City-Ogden	0 800
22	Tulsa	0.798
23	Cincinnati-Hamilton	0.780
24	Nashville	0.752
25	Las Vegas	0.746

Note: NCI: 74.788; -25.0 % change; diameter 11; no disconnects

Table 4-30. Ranking of links based on the removal of St. Louis from the U.S. Internet backbone network

Rank	C/MSA	C/MSA	URCI
1	Chicago-Gary-Kenosha	Washington-Baltimore	0.939
2	Chicago-Gary-Kenosha	Dallas-Fort Worth	0 901
3	Chicago-Gary-Kenosha	Atlanta	0 869
4	Chicago-Gary-Kenosha	New York-Northern New Jersey-Long Island	0.864
5	Chicago-Gary-Kenosha	San Francisco-Oakland-San Jose	0.853
6	Washington-Baltimore	Dallas-Fort Worth	0.850
7	Washington-Baltimore	Atlanta	0.823
8	Washington-Baltimore	New York-Northern New Jersey-Long Island	0.811
9	Washington-Baltimore	San Francisco-Oakland-San Jose	0.798
10	Dallas-Fort Worth	Atlanta	0.792
11	Dallas-Fort Worth	New York-Northern New Jersey-Long Island	0.775
12	Chicago-Gary-Kenosha	Washington-Baltimore	0.769

Table 4-31. U.S. city ranking based on the removal of Sacramento-Yolo from the Internet backbone network

Rank	CMSA	URCI
1	Chicago-Gary-Kenosha	2.873
2	Dallas-Fort Worth	2.597
3	Washington-Baltimore	2 550
4	Atlanta	2.420
5	San Francisco-Oakland-San Jose	2.169
6	New York-Northern New Jersey-Long Island	2.116
7	Kansas City	1.989
8	Los Angeles-Riverside-Orange County	1.635
9	Denver-Boulder-Greeley	1.620
10	St. Louis	1.562
11	Cleveland-Akron	1.305
12	Boston-Worcester-Lawrence	1 296
13	Houston-Galveston-Brazoria	1 285
14	Miami-Fort Lauderdale	1.278
15	Seattle-Tacoma-Bremerton	1.143
16	Indianapolis	1.082
17	Tulsa	0.984
18	Nashville	0.979
19	Pennsauken	0.978
20	Cincinnati-Hamilton	0.872
21	San Diego	0.868
22	Phoenix-Mesa	0.848
23	Charlotte-Gastonia-Rock Hill	0.792
24	Orlando	0.792
25	Tampa-St Petersburg-Clearwater	0.776

Note: NCI: 76.629; -22.9 % change; diameter 11; no disconnects

Table 4-32. Ranking of links based on the removal of Sacramento-Yolo from the U.S. Internet backbone network

Rank	C/MSA	C/MSA	URCI
1	Chicago-Gary-Kenosha	Dallas-Fort Worth	1.046
2	Chicago-Gary-Kenosha	Washington-Baltimore	1.029
3	Chicago-Gary-Kenosha	Atlanta	0 972
4	Dallas-Fort Worth	Washington-Baltimore	0.929
5	Dallas-Fort Worth	Atlanta	0.884
6	Chicago-Gary-Kenosha	San Francisco-Oakland-San Jose	0.877
7	Washington-Baltimore	Atlanta	0.867
8	Chicago-Gary-Kenosha	New York-Northern New Jersey-Long Island	0.857
9	Chicago-Gary-Kenosha	Kansas City	0.803
10	Dallas-Fort Worth	San Francisco-Oakland-San Jose	0 792
11	Washington-Baltimore	San Francisco-Oakland-San Jose	0.775
12	Dallas-Fort Worth	New York-Northern New Jersey-Long Island	0.767

each possible combination of pairs of the top-twelve-ranking nodes in the network, based on URCI. Similar and predictable results where obtained for each of the node pair removal scenarios considered. The results for the pair removal scenarios varied very little from the single-node removal scenarios. For this reason, the results are not explained in the text.

Summary

This chapter has discussed the results of the unweighted analysis of the U.S. Internet Backbone Network. First, the alpha and gamma indices were discussed. A discussion of binary connectivity followed. An explanation of the main network measurements included the URCI, the NCI, and percentage change in connectivity. Other measurements include diameter, disconnects, C/MSA ranking and link ranking. The product of the fully connected network was used to compare changes within the network as each of the top twelve nodes were removed from the network.

The top cities based on connectivity of the U.S. Internet backbone network are (in rank order): Chicago-Gary-Kenosha, Washington-Baltimore, Dallas-Fort Worth, San Francisco-Oakland-San Jose, Atlanta & New York-Northern New Jersey-Long Island. These results are consistent with other Internet infrastructure research. These six C/MSAs lead in ranking of other types of Internet infrastructure as well, including colocation facilities, cellular towers, data centers & web hosting facilities (Townsend 2001, McIntee 2001, Malecki & Gorman 2002, Gorman & McIntee 2003). Telecommunication infrastructure tends to cluster with like kind, creating C/MSAs rich in infrastructure. For example, a C/MSA housing a significant amount of fiber-optic cable will likely have a proportional amount of interconnection facilities to interconnect various networks. Townsend (2001) suggests that the U.S. is experiencing a new type of city, where ranking is based on Internet activity and infrastructure, coining them 'new network cities'. These results confirm Townsend's suggestion that the U.S. is indeed

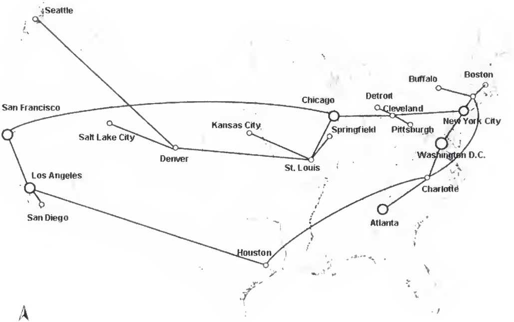


Figure 4-12. Cities connected to the National Science Foundation Network (NSFNET) in 1991(Courtesy of Townsend 2003)

Table 4-33. U.S. city ranking based on the removal of Seattle-Tacoma-Bremerton from the Internet backbone network

Rank	CMSA	URCI
1	Chicago-Gary-Kenosha	2.759
2	Washington-Baltimore	2.695
3	Dallas-Fort Worth	2 656
4	Atlanta	2 448
5	San Francisco-Oakland-San Jose	2.287
6	New York-Northern New Jersey-Long Island	2.185
7	Kansas City	1.927
8	Los Angeles-Riverside-Orange County	1.852
9	Denver-Boulder-Greeley	1.685
10	St. Louis	1.584
11	Sacramento-Yolo	1.393
12	Houston-Galveston-Brazoria	1.373
13	Boston-Worcester-Lawrence	1.328
14	Cleveland-Akron	1.327
15	Miami-Fort Lauderdale	1.306
16	Indianapolis	1.099
17	San Diego	1.056
18	Phoenix-Mesa	1.039
19	Pennsauken	1.019
20	Nashville	1.002
21	Tulsa	0.999
22	Cincinnati-Hamilton	0.889
23	Charlotte-Gastonia-Rock Hill	0.827
24	Orlando	0.815
25	Raliegh-Durham-Chapel Hill	0.803

Note: NCI 80.117; diameter 11; disconnects Anchorage

Table 4-34 Ranking of links based on the removal of Seattle-Tacoma-Bremerton from the U.S. Internet backbone network

Rank	C/MSA	C/MSA	URCI
1	Chicago-Gary-Kenosha	Washington-Baltimore	1.058
2	Chicago-Gary-Kenosha	Dallas-Fort Worth	1.041
3	Washington-Baltimore	Dallas-Fort Worth	1.016
4	Chicago-Gary-Kenosha	Atlanta	0.955
5	Atlanta	Washington-Baltimore	0.935
6	Dallas-Fort Worth	Atlanta	0.922
7	Chicago-Gary-Kenosha	San Francisco-Oakland-San Jose	0.894
8	Washington-Baltimore	San Francisco-Oakland-San Jose	0.871
9	Dallas-Fort Worth	San Francisco-Oakland-San Jose	0.860
10	Chicago-Gary-Kenosha	New York-Northern New Jersey-Long Island	0.855
11	Washington-Baltimore	New York-Northern New Jersey-Long Island	0.832
12	Dallas-Fort Worth	New York-Northern New Jersey-Long Island	0.815

experiencing a new type of city that is dependent on Internet infrastructure. Each type of telecommunication infrastructure measured appears to be growing at a rapid pace in direct proportion to other types of telecommunication infrastructure. In turn, the same cities comprise the top ranking for each of the infrastructures measured

The removal scenarios illustrated the importance of each node removed to the connectivity of the entire network. Each of the nodes removed caused the overall connectivity of the network to drop at least 20%. Fifty percent of the node removal scenarios created at least a 50% decrease in the NCI value. Each of the node removal scenarios had a diameter of 11, except for Denver-Boulder-Greeley & Sacramento-Yolo, their diameter increased to 12. New York-Northern New Jersey-Long Island caused the most disconnects when it was removed from the network, 8.

Centrally located Chicago-Gary-Kenosha ranks highest based on the URCI values of the unweighted U.S. Internet backbone network. Washington-Baltimore, though a coastal city, appears to be nearly as central to the overall network as Chicago-Gary-Kenosha. The Capital ranks second in the network, and causes a significant disturbance to connectivity when removed. This new center of government and high-tech activity is an important node to the unweighted network

The over-simplification of the network (the use of binary weights) contributed greatly to the predictability of the results. The methodology applied with this chapter will be used in the weighted analysis of the U.S. Internet backbone network. Chapter 5 has produced more interesting and relevant findings given that the assumption will be relaxed in favor of a weighted analysis where bandwidth (transmission capacity) is considered. Given that there exists tremendous variability in bandwidth, the results for various node and link removal scenarios will not be obvious.

CHAPTER 5
THE U.S. INTERNET BACKBONE NETWORK: A WEIGHTED ANALYSIS

In this chapter, weighted values are added to the U.S. Internet backbone network. The weighted network is composed of the 218 nodes and 521 links analyzed in Chapter 4. Unlike Chapter 4, however, Chapter 5 incorporates weighted values for the network connections or links. Recall that the network analysis in Chapter 4 relied on a connectivity matrix merely represents the presence or absence of a link using 1's and 0's. By using only binary numbers to represent a link, each of the links would carry equal weight. The analysis in this chapter is more sophisticated in that weights are added to each link. The weights indicate the amount of bandwidth or fiber-optic capacity comprising a link. To review, bandwidth is the term used to describe transmission speed of a fiber-optic cable. The weighted Internet backbone data set provides a measure of the amount of data capacity and connections a city has to move data to another city. The data was calculated from the total long haul fiber capacity, or bandwidth, connecting a C/MSA to other C/MSAs.

A nonbinary matrix was created from the bandwidth data. Like the unweighted analysis, the weighted analysis is based primarily upon matrix multiplication. Before the multiplication process began, the matrix values were converted to lower values to facilitate computation.

Just as the URCI index was modified in Chapter 4 to yield more manageable numbers, so too are the indices scaled in this analysis . The 12 digit numbers were converted into more manageable values by creating a rounded and scaled index. However, unlike the URCI in Chapter 4, this index was created before the matrix

multiplication was performed. Because the weighted values were quite large, the values were converted prior to matrix multiplication due to software limitations. The software would not be able to accommodate the excessively large values. Hence, values in the weighted matrix were divided by the largest number observed in the links for that matrix: 102751.36.

$\text{WRCI}_i = \frac{x_i}{x_{\max}}$ i=1...N where N=47524. When the network was converted to a matrix, there were a total of 47524 cells (218^2), thus N=47524 The largest value in the network (102751.36) occurs at the intersection between Washington-Baltimore and Philadelphia-Wilmington-Atlantic City. The highest value (102751.36) was divided by itself, reassigning the number to a value of 1.0 ; $\text{WRCI}_i = \frac{x_i}{x_{\max}} = \frac{102751.36}{102751.36} = 1$. All other values that were converted resulted in values less than 1.0, as the largest value was reassigned to 1.

As in Chapter 4, the percentage change was calculated by subtracting the NCI (for a given removal scenario) from the original connectivity value. The NCI for the fully connected network is 106726.100.

Bandwidth: An Overview

Table 5-1 illustrates the ranking shift based on bandwidth amongst U.S. cities between 1997-2000 Kellerman (2002, p. 145) explains that among this group, seven have consistently ranked at the top of the list, but in different order: Washington, DC; San Francisco; Chicago; New York, Dallas; Los Angeles and Atlanta The data used in Kellerman's (2002) text has been complied from various sources (Wheeler & O'Kelly 1999, Townsend 2001, Gorman & Malecki 2001, Malecki 2002, O'Kelly & Grubesic 2002). In 1997, Washington, DC ranked first amongst bandwidth, with Chicago and San Francisco following closely behind. In 1998, New York dropped to 5th, down one place from the 1997 measurements. In 1999, New York dropped yet again, ranking 6th in terms of Internet bandwidth. By 2000, however, New York jumped in ranking and overtook

Chicago, Washington, DC, San Francisco, and Dallas as the top node in terms of bandwidth. These cities remained consistently strong in terms of bandwidth measurement throughout the period, with only subtle shifts in their ranking.

Table 5-1. Top-ten U.S. metropolitan areas in total bandwidth on Internet backbones serving them 1997-2000

1997	1998	1999	2000
Washington	San Francisco	Washington	New York
Chicago	Chicago	Dallas	Chicago
San Francisco	Washington	San Francisco	Washington
New York	Dallas	Atlanta	San Francisco
Dallas	New York	Chicago	Dallas
Atlanta	Los Angeles	New York	Atlanta
Los Angeles	Denver	Los Angeles	Los Angeles
Denver	Atlanta	Kansas City	Seattle
Seattle	Seattle	Houston	Denver
Phoenix	Philadelphia	St. Louis	Kansas City

(Source: Malecki 2002. Reproduced from Economic Geography by permission of Clark University as shown in Kellerman, 2002 p 145).
Note: These rankings were figured through the total bandwidth of inter-metropolitan backbones serving a city.

Kellerman (2002) has compiled a table (Table 5-2) based on 1999 Internet backbone data. Normally the total bandwidth of Internet backbones serving a metropolitan area is used to determine city ranking within the network. Kellerman notes the importance of another measure, the number of Internet backbone links to other metropolitan areas (2002 p. 144). In 1999, Washington, DC, ranked first in terms of bandwidth, with 28370 million bits per second (Mbps[1]). This accounted for 7.2% of the bandwidth in the U.S. This bandwidth was housed amongst 233 different links that accounted for 9.6% of total number of links in the U.S. Internet backbone network. San Francisco and Dallas followed close behind Washington, DC, in terms of bandwidth, with 25343 Mbps and 25297 Mbps, respectively. Combined, this was roughly 13% of the nation's total Internet backbone bandwidth. San Francisco also housed a significant

[1]Mbps can be defined as megabits (million bits) per second. A measurement of data transfer rate, or how fast the data can move in one second. For example, 9 Mbps would be a transfer rate of 9 million bits in each second.

percentage of the U.S.'s Internet backbone links. 216 at 9%. Chicago connected to other metro areas with a total of 203 links, roughly 8.4% of the total. The top 10 cities in terms of links comprised nearly 60% of the total number of links in the U.S. Internet backbone network. Fifty percent of the total bandwidth in the U.S. in 1999 was concentrated in the top 10 cities (Table 5-2).

Table 5-2. U.S. City rankings for Internet inter-metropolitan backbones, 1999

Links			Total Bandwidth		
C/MSA	Number	% of U.S.	MSA/CMSA	Mbps	% of U.S.
Washington, DC	233	9.6	Washington, DC	28370	7.2
San Francisco	216	8.9	Dallas	25343	6.4
Chicago	203	8.4	San Francisco	25297	6.4
New York	165	6.8	Atlanta	23861	6.1
Dallas	153	6.3	Chicago	23340	5.9
Atlanta	131	5.4	New York	22232	5.6
Los Angeles	121	5.0	Los Angeles	14868	3.8
Seattle	74	3.1	Kansas City	13525	3.4
Boston	63	2.6	Houston	11522	2.9
Houston	62	2.6	St. Louis	9867	2.6
		58.8			50.3

(Source: Kellerman 2002 p 144; Sources Links. Townsend (2001a); bandwidth: Moss & Townsend (2000))

In 2003, the ranking of cities based on Internet backbones shifts (see Table 5-3). New cities joined the list of top ten cities, and cities that had previously been included in the group from 1997-2000 have been pushed further down the list. Table 5-3 compares the 2000-bandwidth data with the 2003-bandwidth data. The highest-ranking metropolitan area in 2003 is New York-Northern New Jersey-Long Island with 234258 Mbps. In 2000, Washington-Baltimore was ranked first with 816096 Mbps, nearly quadruple the 2000 bandwidth leader, New York-Northern New Jersey-Long Island. It should also be noted that New York-Northern New Jersey-Long Island (234258 Mbps) had a lower bandwidth total in 2000 than does the tenth ranked city in 2003, Los Angeles-Riverside-Orange County (293575 Mbps).

Table 5-3. U.S. city rankings for Internet backbones in 2000 and 2003 based on bandwidth and number of redundant Internet backbone links

C/MSA	2003 Bandwidth Rank	2003 Bandwidth Totals (Mbps)	2000 Rank	2000 Bandwidth Total (Mbps)	Number of Redundant Links 2003
Washington-Baltimore	1	816096	3	208159	245
New York-Northern New Jersey-Long Island	2	727986	1	234258	183
San Francisco-Oakland-San Jose	3	480177	4	201772	147
Dallas-Fort Worth	4	433085	5	183571	142
Atlanta	5	353013	6	149200	105
Chicago-Gary-Kenosha	6	326578	2	221738	141
Cleveland-Akron	7	319967	17	61670	64
Jacksonville	8	300909	37	23952	40
Orlando	9	294506	20	45527	58
Los Angeles-Riverside-Orange County	10	293575	7	140649	88
Tampa-St.Petersburg-Clearwater	11	293331	30	30309	46
Miami-Fort Lauderdale	12	280418	22	42137	55
Philadelphia-Wilmington-Atlantic City	13	258591	14	74167	50
Houston-Galveston-Brazoria	14	250405	12	80482	69
Denver-Boulder-Greeley	15	243960	9	97545	65
Kansas City	16	236797	10	89292	59
Sacramento-Yolo	17	225660	24	40701	46
Boston-Worcester-Lawrence	18	225384	13	75044	59
Richmond-Petersburg	19	222156	32	28194	32
Raliegh-Durham-Chapel Hill	20	197036	50	15731	37
Albany-Schenectady-Troy	21	180774	39	18553	22
Tallahassee	22	169952	44	17684	18
Seattle-Tacoma-Bremerton	23	165955	8	109510	58
Norfolk-Virginia Beach-Newport News	24	160156	42	18198	17
Pittsburgh	25	148020	36	25177	25
Buffalo-Niagara Falls	26	145798	46	17175	18
Salt Lake City-Ogden	27	144393	11	87623	28
Detroit-Ann Arbor-Flint	28	143885	18	53262	24

Table 5-3. Continued

C/MSA	2003 Bandwidth Rank	2003 Bandwidth Totals (Mbps)	2000 Rank	2000 Bandwidth Total (Mbps)	Number of Redundant Links 2003
Portland-Salem	29	141746	16	68174	38
Charlotte-Gastonia-Rock Hill	30	131177	26	35440	29
Toledo	31	130000	56	12753	13
Phoenix-Mesa	32	129210	19	45868	37
St. Louis	33	121791	15	69031	38
San Diego	34	120840	23	42061	32
Greensboro-Winston Salem-High Point	35	120200	47	17128	14

The number of redundant Internet backbone links is important when considering disconnection. This research uses a consolidated figure of bandwidth totals, based on all of the links that connect one metro area to another. This means that the bandwidth data for multiple carriers and their links were used to create a total amount of bandwidth that connects C/MSAs to each other, consolidating redundant links.

If a hacker wanted to completely sever the Internet backbone connections between two cities, the physical location of the fiber-optic cables would have to be determined. For example, to disconnect the Internet backbone connection between New York and Chicago, we'd need to determine the ground routes of the fiber connections between these two cities (assuming there are redundant links, multiple locations would need to be determined). As a side note, according to *Wired* magazine snipping the fiber-optic cables on a corporation's network will land you 5-8 years behind bars (Wired Jan 2004 p. 028).

Table 5-3 shows the top 35 cities in terms of bandwidth totals (2003) and the number of redundant links connected to each C/MSA The top four nodes are Washington-Baltimore with 245 redundant links, New York-Northern New Jersey-Long Island with 183 redundant links, San Francisco-Oakland-San Jose with 147 redundant links, and Dallas-Fort Worth with 142 redundant links. Los Angeles-Riverside-Orange County and San Francisco-Oakland-San Jose have 21 redundant links between them, more than other pairs of C/MSAs in the U.S. (Table 5-4). The close proximity of these two large cities might be the best explanation as to why this particular pair has more redundant links than any other pair of metropolitan areas Houston-Galveston-Brazoria and Dallas-Fort Worth have 19 redundant links between them, as do Washington-Baltimore and Philadelphia-Wilmington-Atlantic City (Table 5-4). Geography can again attribute to the heavy connection redundancy between these cities. Many of the pairs are in close proximity to each other. However, some are not geographically adjacent,

such as Washington-Baltimore and Atlanta (14) (Table 5-4). Geography still plays a role connecting primary nodes in the regional sub-networks The relevance of Chicago-Gary-Kenosha's importance to the Internet backbone network is also evident in terms of redundancy. Chicago-Gary-Kenosha is connected to other cities with significant redundancy: Seattle-Tacoma-Bremerton, St. Louis, Kansas City (9), Cleveland-Akron (10) San Francisco-Oakland-San Jose and New York-Northern New Jersey-Long Island (12) (Table 5-4). Townsend (2001) suggests that Chicago-Gary-Kenosha is not considered in the top ranking of network cities, though established as a global city. The importance of Chicago-Gary-Kenosha in terms of redundant connections though, is evident. The increased number of redundant connections a city has to other metro areas makes it more difficult to eliminate that city from another city, or completely disconnect it from the network. It must be noted though that Chicago-Gary-Kenosha has fallen within the national ranking based on Internet backbone bandwidth (Table 5-3).

Fully Connected Weighted Network

Nodal Ranking

The weighted connectivity matrix was powered up to its diameter and a new ranking of nodes resulted. Figure 5-1 shows the distribution of the U.S. Internet backbone nodes based on the Weighted Relative Connectivity Index (WRCI). A ranking of nodes based on the results was then identified (see Table 5-5). The top 25 nodes in the U.S. Internet backbone network appear to be concentrated in the East (Figure 5-2). Only three of the top 25 nodes are located on the West coast, with a few scattered throughout middle-America (Figure 5-2). The concentration of highly connected nodes in California and on the East Coast is consistent with potential demand. These geographic areas boast the highest concentration of the population and employment opportunities (Figure 5-3). It is of little surprise that Washington-Baltimore and New York-Northern New Jersey-Long Island rank first and second respectively. The third C/MSA in the

Table 5-4. Leading pairs of U.S. cities based on redundant Internet backbone links in 2003

Origin C/MSA	Destination C/MSA	Rank	Number of Redundant Links
Los Angeles-Riverside-Orange County	San Francisco-Oakland-San Jose	1	21
Dallas-Fort Worth	Houston-Galveston-Brazoria	2 tie	19
Washington-Baltimore	Philadelphia-Wilmington-Atlantic City	2 tie	19
Boston-Worcester-Lawrence	New York-Northern New Jersey-Long Island	4	18
Washington-Baltimore	New York-Northern New Jersey-Long Island	5	17
Philadelphia-Wilmington-Atlantic City	New York-Northern New Jersey-Long Island	6	16
Portland-Salem	Seattle-Tacoma-Bremerton	7	15
Atlanta	Washington-Baltimore	8	14
Sacramento-Yolo	San Francisco-Oakland-San Jose	9	13
Jacksonville	Orlando	9	13
Los Angeles-Riverside-Orange County	San Diego	9	13
Chicago-Gary-Kenosha	San Francisco-Oakland-San Jose	12	12
Chicago-Gary-Kenosha	New York-Northern New Jersey-Long Island	12	12
Tampa-St.Petersburg-Clearwater	Miami-Fort Lauderdale	12	12
Washington-Baltimore	Richmond-Petersburg	12	12
Houston-Galveston-Brazoria	Austin-San Marcos	16	11
Tampa-St.Petersburg-Clearwater	Orlando	16	11
Atlanta	Charlotte	18	10
Chicago-Gary-Kenosha	Cleveland-Akron	18	10
Denver-Boulder-Greeley	Salt Lake City-Ogden	18	10
Chicago-Gary-Kenosha	Indianapolis	21	9
San Antonio	Austin-San Marcos	21	9
Seattle-Tacoma-Bremerton	Chicago-Gary-Kenosha	21	9
Chicago-Gary-Kenosha	St. Louis	21	9
St. Louis	Kansas City	21	9
Chicago-Gary-Kenosha	Kansas City	21	9
Miami-Fort Lauderdale	Orlando	21	9

Figure 5-1. U.S. Internet backbone nodes based on weighted relative connectivity index

Figure 5-2. Top 25 nodes in U.S. Internet backbone network based on weighted relative connectivity index

ranking, Philadelphia-Wilmington-Atlantic City, is a newcomer to the usual top-five. It is also surprising to see that five Florida cities are included in the top-ten cities in terms of connectivity, Jacksonville (5th), Orlando (6th), Tallahassee (7th), Tampa-St.Petersburg-Clearwater (8th) and Miami-Fort Lauderdale (9th) (Table 5-5). This is not surprising given the recent expansion and growth of telecommunication infrastructure on the Florida peninsula.

In terms of bandwidth ranking, the Florida cities were not quite as prominent as they are in this new connectivity ranking (Table 5-3). Based on bandwidth ranking, Jacksonville ranked 8th, Orlando 9th, Tampa-St.Petersburg-Clearwater 11th, Miami-Fort Lauderdale 12th and Tallahassee 22nd (Table 5-3). It is interesting to note that Tallahassee made a dramatic rise in the bandwidth ranking, jumping from 44th (2000) to 22nd (2003) in terms of bandwidth. The bandwidth growth in other Florida cities has been significant as well (Table 5-3) Jacksonville ranked 37th in 2000, increasing its bandwidth by over 1200% by 2003. Orlando's bandwidth also grew at a rapid pace, it's growth exceeding 600% from 2000 to 2003 (Table 5-3).

Miami has experienced significant growth in bandwidth and in terms of Internet activity. Miami has become an important hub of Internet activity. Miami is home to the new NAP of the Americas, considered the telecommunications gateway to South America (Orlando Sentinel 2001). The new NAP opened in August of 2000 after much tension about location from carriers, public officials, landowners, and developers (Terremark Worldwide 2000). The surge of Florida cities to importance in the U.S. Internet backbone network contradicts the findings of Malecki (2002). "Malecki's study ranked the top 100 cities in bandwidth availability nationwide in 2000 compared to 1998. It found that the Miami-Ft. Lauderdale area slipped from 19th to 22nd; the Jacksonville area went from 25th to 37th and the Tampa-St. Petersburg area fell from 26th to 30th." (Hoover 2000 UF Press Release http://www.napa.ufl.edu/2000news/bandwidth.htm). In

Figure 5-3. Top 25 nodes in U.S. Internet backbone network compared to census 2001 population density

addition, Malecki's study does not consider the connectivity aspects of Florida cities and its accessibility within the network (considering its direct and indirect linkages).

Table 5-5. Ranking of nodes in the U.S. Internet backbone network

C/MSA	Rank	WRCI
Washington-Baltimore	1	7178.152
New York-Northern New Jersey-Long Island	2	6294.686
Philadelphia-Wilmington-Atlantic City	3	5928.210
Atlanta	4	5444.191
Jacksonville	5	4650.768
Orlando	6	4159.662
Tallahassee	7	3265.773
Tampa-St.Petersburg-Clearwater	8	3124.823
Miami-Fort Lauderdale	9	2943.385
Boston-Worcester-Lawrence	10	2889.174
Richmond-Petersburg	11	2594.082
Dallas-Fort Worth	12	2144.253
Albany-Schenectady-Troy	13	2023.978
Houston-Galveston-Brazoria	14	1963.258
Pittsburgh	15	1946.480
Chicago-Gary-Kenosha	16	1932.045
San Francisco-Oakland-San Jose	17	1855.544
Cleveland-Akron	18	1829.695
Raliegh-Durham-Chapel Hill	19	1534.025
Charlotte-Gastonia-Rock Hill	20	1475.806
Los Angeles-Riverside-Orange County	21	1298.731
Hartford	22	1286.932
New Brunswick	23	1217.261
Kansas City	24	1204.916
Sacramento-Yolo	25	1106.174

Note: NCI 106726.100; diameter 11; no disconnects

Malecki indicated that Florida cities were not growing at a competitive pace with the rest of the cities in the U.S. and attributed the low Florida rankings at the time of his study to it's peripheral location. The current network rankings suggest that Florida is in fact keeping pace with the rest of the nation, climbing the hierarchy (Table 5-3). The geography of Florida may be precisely the reason for Florida's climb up the hierarchy. The new NAP in Miami serves as the 'last stop' for the marine cable-landings carrying data through marine cables to South America. Jacksonville's close proximity to Atlanta and Charlotte make it a perfect hub for interconnection and links the bottom southeastern subnetworks with the subnetworks along the eastern seaboard. As many

long-haul fiber routes travel major interstate routes, and Orlando sits at the crossroads of Florida's Turnpike and Interstate 4, and is in the position to reap the benefits of long-haul fiber destined for Miami. The connectivity of Florida cities in the U.S. Internet backbone network, as well as in bandwidth growth and network connectivity, is readily apparent when one considers the results presented in Tables 5-3 and 5-5.

It must also be noted that Chicago-Gary-Kenosha (16) and San Francisco-Oakland-San Jose (17) have been pushed out of their usual positions amongst the top five cities in the hierarchy (Table 5-5).

The diameter of the network is eleven, meaning the network weighted connectivity matrix was powered ten times before each city was connected by some path to each of the other cities in the network. The NCI of the fully connected network is 106726.100. This value will be used in the single and pair removal scenarios to measure change in connectivity.

Link Ranking

The weighted network multiplication results were also used to identify the most critical links in the network (Table 5-6). The highest values in the powered weighted connectivity matrix identified the most critical links in the network, which is between New York-Northern New Jersey-Long Island and Washington-Baltimore (WRCI 266.775). The second-most prominent link connects Philadelphia-Wilmington-Atlantic City and Washington-Baltimore, the third between Philadelphia-Wilmington-Atlantic City and New York-Northern New Jersey-Long Island. Philadelphia-Wilmington-Atlantic City, New York-Northern New Jersey-Long Island and Washington-Baltimore are the top three nodes in the network. In Chapter 4 it became apparent that the top nodes in the network generally comprised the top link pairs as well The same holds true for the weighted analysis. There also appears to be a relationship between link redundancy (Table 5-4) and the ranking of Internet Backbone links (Table 5-6). The redundancies between New

York-Northern New Jersey-Long Island and Washington-Baltimore are the fifth highest in the U.S., with 17 connections (Table 5-4). This redundancy represents the most important link in the ranking of Internet backbone links in the U.S. (Table 5-6). Philadelphia-Wilmington-Atlantic City and Washington-Baltimore have 19 redundant connections, ranking second nationally (Table 5-4). The combination of these two links is the third most critical in the network ranking (Table 5-6).

Table 5-6. Ranking of Internet backbone links in the U.S

Rank	C/MSA	C/MSA	WRCI
1	New York-Northern New Jersey-Long Island	Washington-Baltimore	266.775
2	Philadelphia-Wilmington-Atlantic City	Washington-Baltimore	242.090
3	Philadelphia-Wilmington-Atlantic City	New York-Northern New Jersey-Long Island	222.118
4	Washington-Baltimore	Atlanta	177.842
5	Orlando	Jacksonville	157.530
6	New York-Northern New Jersey-Long Island	Atlanta	154.797
7	Philadelphia-Wilmington-Atlantic City	Atlanta	153.056
8	Jacksonville	Atlanta	148.417
9	Orlando	Atlanta	143.018
10	Washington-Baltimore	Jacksonville	139.425
11	Jacksonville	Tampa-St.Petersburg-Clearwater	123.439
12	Tampa-St.Petersburg-Clearwater	Orlando	118.236

Node Removal Scenarios

To examine the overall impact on network connectivity nine cities were separately removed from the network. The results of each of the single node removal scenarios reveals some interesting regional patterns and some unpredictable consequences. The pair-removal scenarios were created by using each possible combination for the top six C/MSAs based on bandwidth, and by using each possible combination of the top six C/MSAs based on weighted connectivity. A total of 27 different pairs were removed. The results are divided into separate summary tables for the nine different nodes that were used in the pairings. A summary of the pair-removal scenarios that involved the

particular city is also included. Tables with link and node hierarchies for each city are shown. The results for the weighted analysis are shown for each pair-removal scenario, with summary tables containing NCI, percentage change in connectivity from the original network, diameter, and the number of disconnects for each pair removal scenario. First a summary of the single node scenarios will be discussed. Following the summary is a subsection for each city that will be discussed.

Nine single node removal scenarios were performed The top six nodes in the Internet backbone hierarchy were removed, as were the top six nodes based on bandwidth ranking. Bandwidth is an important characteristic, the measurement of fiber capacity, so those nodes (top six) that ranked highest in bandwidth were also removed. Some cities ranked in the top sixth in both bandwidth and the connectivity hierarchy. A total of nine cities were considered in the removal scenarios, based on bandwidth and connectivity ranking (see Table 5-7).

Malecki and Gorman (2001) predicted that as the Internet grew and redundancy increases that the gap between the network core and network periphery grows (p. 98). At the core six metro areas dominate· Washington, DC; Chicago; Dallas-Fort Worth, San Francisco; New York; and Atlanta. The connectivity results of this research find the same six cities at the top of the hierarchy, confirming the prediction and findings of Malecki and Gorman.

Washington-Baltimore, New York-Northern New Jersey-Long and Atlanta are included in the top-ranks for both bandwidth and connectivity: Washington-Baltimore, New York-Northern New Jersey-Long and Atlanta (see Table 5-8) The network diameter for each of the node removal scenarios did not change from the fully connected network, and remained at 11 for each single node removal scenario. The removal of New York-Northern New Jersey-Long caused more disconnects than the other single node removal scenarios Eight cities were disconnected with the removal of the New York-Northern

Table 5-7. Nodes removed from the weighted connectivity matrix network

Nine cities were considered in the removal scenarios:
Washington-Baltimore *
New York-Northern New Jersey-Long Island *
Philadelphia-Wilmington-Atlantic City
Atlanta *
Jacksonville
Orlando
San Francisco-Oakland-San Jose
Dallas-Fort Worth
Chicago-Gary-Kenosha
The top-six cities based on bandwidth measurement:
Washington-Baltimore
New York-Northern New Jersey-Long
San Francisco-Oakland-San Jose
Dallas-Fort Worth
Atlanta *
Chicago-Gary-Kenosha
The top-six cities based on connectivity:
Washington-Baltimore *
New York-Northern New Jersey-Long Island *
Philadelphia-Wilmington-Atlantic City
Atlanta *
Jacksonville
Orlando

Note: * Indicates this node was ranked within the top-six cities for both bandwidth and connectivity

New Jersey-Long. Four disconnects occurred when Chicago-Gary-Kenosha was removed from the network, three with the absence of San Francisco-Oakland-San Jose, and one each for Orlando and Dallas-Fort Worth. The removal of Washington-Baltimore caused the greatest change in the network, with a negative 70% drop in connectivity. The NCI dropped from 106726.1 (fully connected network) to 32227.18. New York-Northern New Jersey-Long Island was the next-most- disruptive removal scenario. Her removal caused 61% change in connectivity and the NCI dropped to 41258.048. Centrally located Atlanta ranks third for the disruption of her removal, with 60% change in connectivity. Atlanta ranks fifth in terms of bandwidth and fourth in connectivity. Jacksonville, the most surprising addition to the usual hierarchy, ranks fifth in terms of

Table 5-8. Summary of single node removal scenarios from the U.S. Internet backbone network

	Connectivity Rank	Bandwidth Rank	Disruption Rank	Network Connectivity Index (NCI)	% Change	Diameter	Disconnects
Initial Network Values:				106726.100	N/A	11	N/A
C/MSAs Removal Scenarios:							
Washington-Baltimore	1	1	1	32227.180	-69.8	11	0
New York-Northern New Jersey-Long Island	2	2	2	41258.048	-61.3	11	8
Philadelphia-Wilmington-Atlantic City	3	13	4	49301.577	-53.8	11	0
Atlanta	4	5	3	43549.641	-59.1	11	0
Jacksonville	5	8	5	65544.805	-38.5	11	0
Orlando	6	9	6	73100.024	-31.5	11	1
San Francisco-Oakland-San Jose	17	3	9	92284.869	-13.5	11	3
Dallas-Fort Worth	12	4	7	87348.944	-18.1	11	1
Chicago-Gary-Kenosha	16	6	8	89811.660	-15.8	11	4

disruption and connectivity, and eighth for bandwidth. The removal of Jacksonville causes a 39% change in connectivity. Orlando's removal causes nearly the same amount of change in connectivity as the removal of Jacksonville. The connectivity changes 32% with the absence of Orlando. Dallas-Fort Worth (19%), Chicago-Gary-Kenosha (16%) and San Francisco-Oakland-San Jose (14%) caused the least amount of connectivity disruption to the network upon their removals.

Washington-Baltimore

Washington-Baltimore ranks first in both bandwidth and in terms of connectivity for the U.S. Internet backbone network. It comes as little surprise that the removal of the highest-ranking city in these measurements would cause the most disruption to the network. With the removal of this city, it is a surprise to see Orlando rising to the top of the ranking of highly connected cities (Table 5-9). Two other Florida-cities follow in second and third place, Jacksonville and Tampa-St.Petersburg-Clearwater. Atlanta garners fourth rank. Two more Sunshine state cities rank fifth and sixth; Miami-Fort Lauderdale and Tallahassee (Table 5-9). The removal disconnected no other C/MSAs from the network. However, the connectivity index decreased more than in any other removal scenario confirming the importance of Washington-Baltimore to the overall network (NCI-32227.180).

The results in Table 5-10 highlight the new importance of Florida in the ranking of links is even more prominent. In fact, the top-five links are pairs of Florida cities. Atlanta moves into the sixth and seventh rank with Orlando and then Jacksonville. The remaining top-twelve links are also pairs of Florida cities (Table 5-9).

The pair-removal scenarios involving Washington-Baltimore are shown in Table 5-11. The pairs had fairly consistent results with each other; the average change in network connectivity was between -75% and -90%. The removal of Washington-Baltimore with Philadelphia-Wilmington-Atlantic City caused a significant drop in connectivity, the NCI drops 72%. The close proximity of these two nodes leads one to

believe they are well connected, thus causing a more significant impact upon their removal. However, their geography seems to have had the opposite effect. It appears they are less dependent upon each other than the other pairs. This could attributed to Philadelphia-Wilmington-Atlantic City's lower ranking in bandwidth (Table 5-3).

Table 5-9. U.S. city rankings based on the removal of Washington-Baltimore from the Internet backbone network

Rank	C/MSA	WRCI
1	Orlando	2566.213
2	Jacksonville	2525.438
3	Tampa-St.Petersburg-Clearwater	2004.688
4	Atlanta	1872.215
5	Miami-Fort Lauderdale	1810.133
6	Tallahassee	1724.544
7	Houston-Galveston-Brazoria	922.037
8	Dallas-Fort Worth	892.226
9	New York-Northern New Jersey-Long Island	771.827
10	Daytona Beach	582.129
11	San Francisco-Oakland-San Jose	573.166
12	Chicago-Gary-Kenosha	554 434
13	Kansas City	519.556
14	Boston-Worcester-Lawrence	514.444
15	Fort Myers-Cape Coral	501.532
16	Los Angeles-Riverside-Orange County	496.585
17	Denver-Boulder-Greeley	473.741
18	West Palm Beach-Boca Raton	456.521
19	Sacramento-Yolo	431.501
20	Charlotte-Gastonia-Rock Hill	395.260
21	Lakeland-Winterhaven	373.315
22	Cleveland-Akron	346.141
23	New Orleans	341.702
24	Philadelphia-Wilmington-Atlantic City	331.655
25	San Diego	328.546

Note: NCI 32227.180; -69.8 % change; diameter 11; no disconnects

The Washington-Baltimore node has consistently ranked near the top of the ranking of U S. cities for bandwidth measurement, redundant backbone links (Table 5-3), fiber-lit buildings and colocation facilities (Table 4-5). Kellerman notes that in 1999 Washington-Baltimore ranked first in both number of links and total bandwidth but that the dominance of the city in the number of links was higher than in total bandwidth. He attributed this unbalance as testimonial to the "central role as a capital city" (Kellerman 2002, p. 144). The capital is home to one of the nation's largest interconnection and peering points; MAE East (housed in two facilities MAE-East and MAE-East+).

According to Kellerman (2002, p. 93) the largest number of Internet Service Providers (ISP), 92, connect in MAE-East. The following other top international connection points are Chicago NAP (83), MAE-West San Jose (83) and LINX (The London Internet Exchange) (82) (Kellerman 2002, p. 93). The coastal position of Washington-Baltimore has helped push the city to international prominence for Internet backbones and interconnection. This city ranked seventh in the world for intercontinental backbones in 2001 (Kellerman 2002 p 147).

Table 5-10. Ranking of links based on the removal of Washington-Baltimore from the U.S. Internet backbone network

Rank	C/MSA	C/MSA	WRIC
1	Orlando	Jacksonville	269.63
2	Orlando	Tampa-St.Petersburg-Clearwater	215.00
3	Jacksonville	Tampa-St.Petersburg-Clearwater	206.47
4	Miami-Fort Lauderdale	Orlando	192.92
5	Jacksonville	Miami-Fort Lauderdale	190.25
6	Orlando	Atlanta	187.76
7	Jacksonville	Atlanta	186.92
8	Orlando	Tallahassee	182.79
9	Jacksonville	Tallahassee	177.90
10	Miami-Fort Lauderdale	Tampa-St.Petersburg-Clearwater	151.81
11	Tampa-St.Petersburg-Clearwater	Atlanta	148.55
12	Tampa-St.Petersburg-Clearwater	Tallahassee	141.76

Table 5-11. Summary of node removal pairs with Washington-Baltimore from the U.S. Internet backbone network

	Network Connectivity Index (NCI)	% Change	Diameter	Disconnects
Initial network values	106726.1	N/A	11	N/A
Double node removal scenarios paired with Washington-Baltimore				
Orlando	11589.850	-89.1	11	1
New York-Northern New Jersey-Long Island	24792.020	-76.8	11	8
Philadelphia-Wilmington-Atlantic City	29687.770	-72.1	11	0
Atlanta	13488.598	-87.5	11	0
Jacksonville	11389 702	-89 3	11	0
San Francisco-Oakland-San Jose	26449.206	-75.2	11	3
Dallas-Fort Worth	22294 329	-79.1	11	1
Chicago-Gary-Kenosha	25592.043	-76.0	11	4

New York-Northern New Jersey-Long Island

Like Washington-Baltimore, the New York-Northern New Jersey-Long Island hub is an important international player in the Internet backbone and interconnection game. The international lines with the largest bandwidth capacity in 2000 connected London and New York as well as London and Paris (Telegeography 2001, 2000). New York also leads in intercontinental backbones (Kellerman 2002, p. 147), partially attributed to it being one of the original peering points established by the National Science Foundation in 1994. This helped to establish NYC as an interconnection hub for Internet backbones on a domestic level, and then later on the global level. Although New York lagged in bandwidth between 1997-1999, it moved to the top of the ranking by 2000. However, by 2003 New York had already been pushed back to second rank in bandwidth (Table 5-8). This may also explain why New York ranks second in terms of connectivity (Table 5-8), and second in terms of Web development firms (Malecki 2000). New York-Northern New Jersey-Long Island was ranked second by Zook (2000) for domain names[2] as well. Second place appears to be a developing pattern for the Big Apple. It should also be noted that 75% of U.S. e-commerce is estimated to run through five websites: Amazon.com; ebay; AOL; Yahoo!; and Buy.com (Button & Taylor 2001), with AOL located in New York.

With the removal of New York-Northern New Jersey-Long Island, we see a significant impact upon the connectivity of the network. The removal ranks second in terms of disturbance upon the network. The NCI drops 61% to 41258.048. Southern cities, particularly those located in Florida and the Southwest are again high in the new

[2]Domain names are the unique name that identifies an Internet site, they always have 2 or more parts, separated by dots. The part on the left is the most specific, and the part on the right is the most general. A given machine may have more than one Domain Name but a given Domain Name points to only one machine.

hierarchy of C/MSA's. Jacksonville rises to the top of the ranking, followed by Orlando (Table 5-12). Atlanta is bumped up to the third rank. The next three ranks are again filled by Florida cities: Tampa-St.Petersburg-Clearwater, Miami-Fort Lauderdale and Tallahassee. Nine of the top-twelve-ranking cities are geographically located in the South, seven in the Southeast. San Francisco-Oakland-San Jose, ranking 13th, is the first West coast city included in the ranking. Atlanta has also gained rank within the ranking of links (Table 5-13). The connection between Jacksonville and Orlando is the most important link, consistent with the first and second ranking they hold within the ranking of nodes (Table 5-12). The next two most critical links are between Atlanta and Jacksonville, then Atlanta and Orlando.

Table 5-12. U.S. city ranking based on the removal of New York-Northern New Jersey-Long Island from the Internet backbone network

Rank	C/MSA	WRCI
1	Jacksonville	3172.118
2	Orlando	3118.694
3	Atlanta	2669.820
4	Tampa-St.Petersburg-Clearwater	2402.355
5	Miami-Fort Lauderdale	2191.098
6	Tallahassee	2170.274
7	Washington-Baltimore	1724.463
8	Houston-Galveston-Brazoria	1174.601
9	Dallas-Fort Worth	1133.706
10	Philadelphia-Wilmington-Atlantic City	966.834
11	Richmond-Petersburg	699.782
12	Daytona Beach	699.054
13	San Francisco-Oakland-San Jose	668.415
14	Charlotte-Gastonia-Rock Hill	662.782
15	Los Angeles-Riverside-Orange County	595.898
16	Fort Myers-Cape Coral	591.844
17	Raliegh-Durham-Chapel Hill	571.418
18	West Palm Beach-Boca Raton	548.919
19	Kansas City	546.474
20	Denver-Boulder-Greeley	511.697
21	Pittsburgh	459.676
22	Sacramento-Yolo	459.274
23	Lakeland-Winterhaven	443.356
24	Chicago-Gary-Kenosha	441.972
25	New Orleans	413.453

Note: NCI 41258.0481; -61.3 % change; diameter 11; disconnects: Wayne, Hackensack, Freehold, Cedar Knolls, Rochelle Park, Bohemia, Garden City, Bridgeport

In the pair removal summary table (Table 5-14) Philadelphia-Wilmington-Atlantic City as a pair partner causes the least disruption. The change in network connectivity changes only 37% with the removal of New York-Northern New Jersey-Long Island and Philadelphia-Wilmington-Atlantic City. The NCI of this pair drops to 67023 276. Their removal does disconnect 10 C/MSAs.

Table 5-13. Ranking of links based on the removal of New York-Northern New Jersey-Long Island from the U.S. Internet backbone network

Rank	C/MSA	C/MSA	WRIC
1	Orlando	Jacksonville	295.905
2	Jacksonville	Atlanta	238.022
3	Orlando	Atlanta	228.329
4	Orlando	Tampa-St.Petersburg-Clearwater	227.908
5	Jacksonville	Tampa-St.Petersburg-Clearwater	223.950
6	Jacksonville	Miami-Fort Lauderdale	207.716
7	Orlando	Miami-Fort Lauderdale	205.864
8	Orlando	Tallahassee	200.796
9	Tallahassee	Jacksonville	199.780
10	Tampa-St.Petersburg-Clearwater	Atlanta	179.382
11	Miami-Fort Lauderdale	Atlanta	162.367
12	Tallahassee	Atlanta	161.225

Table 5-14. Summary of node removal pairs with New York-Northern New Jersey-Long Island from the U S. Internet backbone network

	Network Connectivity Index (NCI)	% Change	Diameter	Disconnects
Initial Network Values	106726.1	N/A	11	N/A
Double node removal scenarios				
C/MSAs removed with New York-Northern New Jersey-Long Island:				
Philadelphia-Wilmington-Atlantic City	67023.276	-37.2	11	10
Atlanta	12754.210	-88 0	11	8
Jacksonville	13438.667	-87.4	11	8
Orlando	15068.828	-85.9	11	9
Washington-Baltimore	24792.020	-76.8	11	8
Dallas-Fort Worth	28376.720	-73.4	11	9
Chicago-Gary-Kenosha	36665.223	-65.6	11	12
San Francisco-Oakland-San Jose	34622.124	-67.6	11	11

Philadelphia-Wilmington-Atlantic City

Due to Philadelphia-Wilmington-Atlantic City low ranking in bandwidth (13), it is surprising that this C/MSA ranks third in the ranking of connectivity and fourth in terms of

disruption. Philadelphia-Wilmington-Atlantic City made the top ten ranking for bandwidth only once between 1997 and 2000 (Table 5-1). This node ranked 10th in 1998 and ranked 14th in 2000. Ranking improved to 13th by 2003, but was still low. Philadelphia-Wilmington-Atlantic City is not included within the top 10 ranks of other types of Internet infrastructure measurements including bandwidth, fiber lit buildings, and colocation facilities. Philadelphia-Wilmington-Atlantic City does seem to be gaining importance is in the wireless industry. This C/MSA ranked sixth in the U.S., housing 960 cell towers in 2001 (Gorman & McIntee 2003). This surge in wireless infrastructure and lack of fixed infrastructure suggest that Philadelphia-Wilmington-Atlantic City is moving toward wireless data transfer.

Table 5-15. U.S. city rankings based on the removal of Philadelphia-Wilmington-Atlantic City from the Internet backbone network

Rank	C/MSA	WRCI
1	Jacksonville	3416.146
2	Orlando	3334.204
3	Atlanta	2971.499
4	Tampa-St.Petersburg-Clearwater	2557.047
5	Tallahassee	2339.470
6	Miami-Fort Lauderdale	2335.742
7	Washington-Baltimore	1677.640
8	New York-Northern New Jersey-Long Island	1428.480
9	Dallas-Fort Worth	1291.929
10	Houston-Galveston-Brazoria	1288.900
11	San Francisco-Oakland-San Jose	893.515
12	Boston-Worcester-Lawrence	848.755
13	Chicago-Gary-Kenosha	817.785
14	Daytona Beach	744.120
15	Los Angeles-Riverside-Orange County	728.740
16	Charlotte-Gastonia-Rock Hill	717.170
17	Kansas City	713.209
18	Richmond-Petersburg	687.945
19	Denver-Boulder-Greeley	639.656
20	Cleveland-Akron	636.913
21	Fort Myers-Cape Coral	626.718
22	Sacramento-Yolo	611.308
23	Raliegh-Durham-Chapel Hill	594.307
24	West Palm Beach-Boca Raton	584.164
25	Albany-Schenectady-Troy	523.321

Note: NCI 49301.577; -53.8 % change; diameter 11; no disconnects

The ranking of nodes when Philadelphia-Wilmington-Atlantic City is removed looks very similar to the ranking with the removal of New York-Northern New Jersey-Long Island. Cities like Jacksonville, Orlando, Atlanta, Tampa-St.Petersburg-Clearwater, and Tallahassee climb to the top of the ranking (Table 5-16). Table 5-17 shows only pairs of Florida cities, and Atlanta paired with Florida cities.

Table 5-16. Ranking of links based on the removal of Philadelphia-Wilmington-Atlantic City from the U.S. Internet backbone network

Rank	C/MSA	C/MSA	WRIC
1	Orlando	Jacksonville	299.897
2	Atlanta	Jacksonville	243.798
3	Atlanta	Orlando	236.153
4	Tampa-St.Petersburg-Clearwater	Orlando	230.161
5	Tampa-St.Petersburg-Clearwater	Jacksonville	227.041
6	Jacksonville	Miami-Fort Lauderdale	210.583
7	Miami-Fort Lauderdale	Orlando	208.068
8	Jacksonville	Tallahassee	203.906
9	Tallahassee	Orlando	203.672
10	Atlanta	Tampa-St Petersburg-Clearwater	184.206
11	Atlanta	Miami-Fort Lauderdale	167.061
12	Atlanta	Tallahassee	165.765

Table 5-17. Summary of node removal pairs with Philadelphia-Wilmington-Atlantic City from the U.S. Internet backbone network

	Network Connectivity Index (NCI)	% Change	Diameter	Disconnects
Initial Network Values	106726.1	N/A	11	N/A
Double node removal scenarios				
C/MSAs removed with Philadelphia-Wilmington-Atlantic City:				
Atlanta	15969.913	-85.0	11	0
Jacksonville	18477.761	-82.7	11	0
Orlando	20814.592	-80.5	11	1
New York-Northern New Jersey-Long Island	67023.276	-37.2	11	10
Washington-Baltimore	29687.770	-72.1	11	0

As was mentioned earlier in the text, the pairing of Philadelphia-Wilmington-Atlantic City and New York-Northern New Jersey-Long Island causes only a -37% change in the NCI. The removal of Philadelphia-Wilmington-Atlantic City and Atlanta caused the greatest connectivity decline of the Philadelphia-Wilmington-Atlantic City

pair-scenarios, ~85%. It is important to note that New York-Northern New Jersey-Long Island and Washington-Baltimore are geographically closest to Philadelphia-Wilmington-Atlantic City. This is most likely attributed to high redundant connections within the Bo-Wash corridor, thus causing less disruption to the overall network when a node within the region is removed.

Atlanta

Atlanta's local government plays an active role in the development and implementation of telecommunication infrastructure into its city. This benefits both telecommunication firms and the city. In 1999 Atlanta contained operations for 431 of Fortune's top 500 industrial firms in the U.S., and this may be partly attributed to her sophisticated infrastructure (Kotval 1999). Atlanta has maintained top-ten ranking in bandwidth since 1997 (see Table 5-1 and 5-3). In 2003 Atlanta ranked fifth in bandwidth.

When Atlanta is removed from the network, the ranking of nodes is radically different from the rankings so far discussed (Table 5-18). There is no sign of the Florida C/MSAs that dominate the ranking with Atlanta's presence. This leads us to conclude that much of the Florida connectivity to the rest of the backbone network is highly dependent upon interconnection through Atlanta, making it the critical hub. The ranking looks much like the original network hierarchy with Washington-Baltimore, New York-Northern New Jersey-Long Island, and Philadelphia-Wilmington-Atlantic City ranking first through third respectively (Table 5-18). For the first time we see Boston join the top-five, ranking fourth in this new hierarchy. The removal of Atlanta causes the third-highest absolute change in connectivity (Table 5-8).

As previously mentioned, the ranking of links show a notable absence of Florida cities (Table 5-19). The most critical link is between Washington-Baltimore and New York-Northern New Jersey-Long Island. This reiterates the pattern of the connection between the first-and second-ranked city to be the most important link. One of the most

prominent pair removal scenario with Atlanta is Chicago-Gary-Kenosha (Table 5-20). The NCI drops 70.9%. These cities are both centrally located, though one in the north and one in the south. The simultaneous removal of these two central nodes has a radical effect upon the network. Both Atlanta & Chicago-Gary-Kenosha are dominant hubs in their regions.

Table 5-18. U.S. city rankings based on the removal of Atlanta from the Internet backbone network

Rank	C/MSA	WRCI
1	Washington-Baltimore	3390.104
2	New York-Northern New Jersey-Long Island	3351.031
3	Philadelphia-Wilmington-Atlantic City	3137.567
4	Boston-Worcester-Lawrence	1608.950
5	Richmond-Petersburg	1233.497
6	Chicago-Gary-Kenosha	1145.671
7	Cleveland-Akron	1133.889
8	San Francisco-Oakland-San Jose	1131.259
9	Albany-Schenectady-Troy	1130.441
10	Pittsburgh	1037.813
11	Orlando	873.695
12	Dallas-Fort Worth	855.994
13	Tampa-St.Petersburg-Clearwater	779.974
14	Los Angeles-Riverside-Orange County	757.158
15	Sacramento-Yolo	727.482
16	Kansas City	710.704
17	Hartford	706.780
18	Jacksonville	660.764
19	Miami-Fort Lauderdale	658.355
20	New Brunswick	653.241
21	Denver-Boulder-Greeley	614.816
22	Houston-Galveston-Brazoria	574.652
23	Providence-Fall River-Warwick	565.719
24	Harrisburg-Lebanon-Carlisle	532.902
25	Seattle-Tacoma-Bremerton	518.867

Note: NCI 43549.641; -59.1 % change; diameter 11; no disconnects

Jacksonville

Jacksonville ranks fifth in both connectivity and the percentage of disruption amongst the node removals (Table 5-8). Jacksonville is a new addition to the Internet rankings. In Table 5-3 it was apparent that Jacksonville has been growing tremendously in terms of bandwidth. Jacksonville ranked 37th in 2000, but had boosted to 8th by 2003.

It seems that the infrastructure building in Jacksonville will continue. *Business Florida* reports that a NAP is planned for Jacksonville and that it will provide better service between Florida and Europe (Bleyer 2004). With a future NAP in Jacksonville and two existing NAPs in Miami, the Florida peninsula will continue to experience a surge in telecommunication infrastructure builds, particularly Internet backbone fiber.

Table 5-19. Ranking of links based on the removal of Atlanta from the U.S. Internet backbone network

Rank	C/MSA	C/MSA	WRIC
1	Washington-Baltimore	New York-Northern New Jersey-Long Island	323.030
2	Washington-Baltimore	Philadelphia-Wilmington-Atlantic City	312.326
3	New York-Northern New Jersey-Long Island	Philadelphia-Wilmington-Atlantic City	303.897
4	New York-Northern New Jersey-Long Island	Boston-Worcester-Lawrence	154.176
5	Washington-Baltimore	Boston-Worcester-Lawrence	152.584
6	Philadelphia-Wilmington-Atlantic City	Boston-Worcester-Lawrence	141.154
7	Washington-Baltimore	Richmond-Petersburg	125.849
8	Richmond-Petersburg	New York-Northern New Jersey-Long Island	118.670
9	Albany-Schenectady-Troy	New York-Northern New Jersey-Long Island	111.790
10	Richmond-Petersburg	Philadelphia-Wilmington-Atlantic City	110.006
11	Albany-Schenectady-Troy	Washington-Baltimore	109 371
12	Cleveland-Akron	Washington-Baltimore	105.820

Table 5-20. Summary of node removal pairs with Atlanta from the U.S. Internet backbone network

	Network Connectivity Index (NCI)	% Change	Diameter	Disconnects
Initial Network Values	106726.1	N/A	11	N/A
Double node removal scenarios C/MSAs removed with Atlanta:				
Jacksonville	39117.942	-63.3	11	1
Orlando	38017.184	-64.3	11	1
Philadelphia-Wilmington-Atlantic City	15969.913	-85.0	11	0
New York-Northern New Jersey-Long Island	12754.210	-88.0	11	8
Washington-Baltimore	13488.598	-87.3	11	0
Chicago-Gary-Kenosha	31044 025	-70 9	11	4
Dallas-Fort Worth	34543.589	-67.6	11	1
San Francisco-Oakland-San Jose	32276.762	-69.7	11	3

The ranking shown in Table 5-21 is very similar to the original network ranking prior to the removal of Jacksonville. The absence of Jacksonville, however there is also a notable absence of other Florida cities in the ranking. Tallahassee is the only Florida city in the top 25, squeezing into the 25th rank. In Table 5-22 Richmond-Petersburg joins the group of top 12 links for the first time in its connection between New York-Northern New Jersey-Long Island and Philadelphia-Wilmington-Atlantic City.

Table 5-21. U.S. city rankings based on the removal of Jacksonville from the Internet backbone network

Rank	C/MSA	WRCI
1	Washington-Baltimore	5713.504
2	New York-Northern New Jersey-Long Island	5250.261
3	Philadelphia-Wilmington-Atlantic City	4931.353
4	Atlanta	2430.026
5	Boston-Worcester-Lawrence	2395.030
6	Richmond-Petersburg	2016.620
7	Albany-Schenectady-Troy	1676.184
8	Chicago-Gary-Kenosha	1600.493
9	Cleveland-Akron	1581.112
10	Pittsburgh	1578.169
11	San Francisco-Oakland-San Jose	1555.199
12	Dallas-Fort Worth	1359.107
13	Hartford	1053.369
14	Los Angeles-Riverside-Orange County	1046.615
15	Houston-Galveston-Brazoria	1020.335
16	Raliegh-Durham-Chapel Hill	1001.382
17	New Brunswick	988.137
18	Kansas City	959.190
19	Sacramento-Yolo	952.753
20	Harrisburg-Lebanon-Carlisle	842.402
21	Providence-Fall River-Warwick	823.460
22	Denver-Boulder-Greeley	814.762
23	Charlotte-Gastonia-Rock Hill	798.949
24	Charlottesville	769.888
25	Tallahassee	732.759

Note: NCI 65544.805; -38.5 % change; diameter 11; no disconnects

Orlando

In 2000, the city of Orlando ranked 20th in bandwidth, and by 2003 had moved to the ninth place in the ranking (Table 5-3). Orlando ranked sixth in the connectivity of the original network, and was the sixth most disruptive removal scenario of the nine considered (Table 5-8). Like Jacksonville, Orlando has experienced rapid growth in

Table 5-22. Ranking of links based on the removal of Jacksonville from the U.S. Internet backbone network

Rank	C/MSA	C/MSA	WRIC
1	Washington-Baltimore	New York-Northern New Jersey-Long Island	486.07
2	Washington-Baltimore	Philadelphia-Wilmington-Atlantic City	473.28
3	New York-Northern New Jersey-Long Island	Philadelphia-Wilmington-Atlantic City	427.56
4	Washington-Baltimore	Atlanta	226.26
5	Washington-Baltimore	Boston-Worcester-Lawrence	220.82
6	New York-Northern New Jersey-Long Island	Boston-Worcester-Lawrence	206.64
7	New York-Northern New Jersey-Long Island	Atlanta	202.16
8	Washington-Baltimore	Richmond-Petersburg	197.51
9	Philadelphia-Wilmington-Atlantic City	Boston-Worcester-Lawrence	190.07
10	Philadelphia-Wilmington-Atlantic City	Atlanta	185.45
11	New York-Northern New Jersey-Long Island	Richmond-Petersburg	173.87
12	Philadelphia-Wilmington-Atlantic City	Richmond-Petersburg	159.03

Table 5-23. Summary of node removal pairs with Jacksonville from the U.S. Internet backbone network

	Network Connectivity Index (NCI)	% Change	Diameter	Disconnects
Initial Network Values	106726.1	N/A	11	N/A
Double node removal scenarios C/MSAs Removed with Jacksonville:				
Philadelphia-Wilmington-Atlantic City	18477.761	-82.7	11	0
Atlanta	39117.942	-63.3	11	1
Orlando	60812.004	-43.0	11	1
New York-Northern New Jersey-Long Island	13438 667	-87.4	11	8
Washington-Baltimore	11389.702	-89.3	11	0

Internet bandwidth and other types of telecommunication infrastructure (colocation, digital telephone switches, wireless structure). An Internet consumption comparison of U.S. cities in 2001 ranked Orlando eighth for Internet penetration in the home. The consumption study found that 61.4% of the Orlando market had experienced the Internet in their home as regular users (Internet com 2001, Nielsen//NetRatings). The geographical location of Orlando as a midway point between Miami and Jacksonville can only be viewed as a contributing factor in Orlando's rise in ranking of Internet activity.

When Orlando is removed, Florida cities are pushed down the ranking Note that Jacksonville, ranking 12th, is the first Florida city listed in the ranking. Tallahassee holds the 18th rank and Miami is not included in the top 25 nodes of the ranking (Table 5-24). This confirms the theory that Orlando is an important hub of interconnection for Miami to the rest of the Internet backbone No Florida cities are included in the ranking of links when Orlando is removed from the network (Table 5-24). As with the removal of the other Florida city, Jacksonville, we see Boston rise in the ranking of both links and nodes (Tables 5-24, 5-25). The network experiences the most connectivity change when Jacksonville and Washington-Baltimore are removed together, 89%. This is unsurprising, as Jacksonville and Washington-Baltimore are two predominant hubs in their regional subnetworks.

Table 5-24. U.S. city rankings based on the removal of Orlando from the Internet backbone network

Rank	C/MSA	WRCI
1	Washington-Baltimore	6322.338
2	New York-Northern New Jersey-Long Island	5727.533
3	Philadelphia-Wilmington-Atlantic City	5381.214
4	Atlanta	3237.451
5	Boston-Worcester-Lawrence	2585.293
6	Richmond-Petersburg	2217.650
7	Albany-Schenectady-Troy	1807.970
8	Pittsburgh	1711.240
9	Chicago-Gary-Kenosha	1699.575
10	Cleveland-Akron	1683.688
11	San Francisco-Oakland-San Jose	1646.399
12	Jacksonville	1524.387
13	Dallas-Fort Worth	1452 666
14	Raliegh-Durham-Chapel Hill	1153.087
15	Hartford	1137.232
16	Houston-Galveston-Brazoria	1119.939
17	Los Angeles-Riverside-Orange County	1107.234
18	Tallahassee	1103.459
19	New Brunswick	1069.673
20	Kansas City	1003.151
21	Sacramento-Yolo	997.281
22	Charlotte-Gastonia-Rock Hill	980.790
23	Harrisburg-Lebanon-Carlisle	920.110
24	Providence-Fall River-Warwick	884.643
25	Denver-Boulder-Greeley	848.269

Note. NCI 73100.024; -31.5 % change; diameter 11; disconnects Ojus

Table 5-25. Ranking of links based on the removal of Orlando from the U.S. Internet backbone network

Rank	C/MSA	C/MSA	WRIC
1	Washington-Baltimore	New York-Northern New Jersey-Long Island	517.132
2	Washington-Baltimore	Philadelphia-Wilmington-Atlantic City	504.086
3	New York-Northern New Jersey-Long Island	Philadelphia-Wilmington-Atlantic City	448.784
4	Atlanta	Washington-Baltimore	297.427
5	Atlanta	New York-Northern New Jersey-Long Island	260.315
6	Atlanta	Philadelphia-Wilmington-Atlantic City	238.393
7	Washington-Baltimore	Boston-Worcester-Lawrence	232.924
8	New York-Northern New Jersey-Long Island	Boston-Worcester-Lawrence	215.037
9	Washington-Baltimore	Richmond-Petersburg	211.864
10	Boston-Worcester-Lawrence	Philadelphia-Wilmington-Atlantic City	197.684
11	New York-Northern New Jersey-Long Island	Richmond-Petersburg	183.793
12	Richmond-Petersburg	Philadelphia-Wilmington-Atlantic City	167.525

Table 5-26. Summary of node removal pairs with Orlando from the U.S. Internet backbone network

	Network Connectivity Index (NCI)	% Change	Diameter	Disconnects
Initial Network Values	106726.1	N/A	11	N/A
Double node removal scenarios C/MSAs removed with Orlando:				
Jacksonville	60812.004	-43.0	11	1
Atlanta	38017.184	-64.4	11	1
Philadelphia-Wilmington-Atlantic City	20814.592	-80.5	11	1
New York-Northern New Jersey-Long Island	15068.828	-85.9	11	9
Washington-Baltimore	11589.850	-89.1	11	1

San Francisco-Oakland-San Jose

San Francisco-Oakland-San Jose is the first West-coast hub to be removed from the network. The San Francisco-Oakland-San Jose area ranks third in bandwidth nationally but only 17th in terms of connectivity (Table 5-8). Though San Francisco-Oakland-San Jose does not rank well in connectivity, it does rank well in other types of Internet related infrastructure. The San Francisco-Oakland-San Jose CMSA ranks third in fiber-lit buildings (GeoTel 2003), cellular structure (Gorman & McIntee 2003), and

colocation facilities (McIntee 2001); and fifth in domain names (Zook 2000). San Francisco-Oakland-San Jose has been an important Internet hub since the beginning of the Internet; from ARPANET through NSFNET to the Internet (Kellerman 2002, p. 141). There is a clustering of Internet backbone interconnection facilities is in the San Francisco C/MSA, connecting the high concentration of telecommunication networks. One of the four main NAPs established by the NSF is located in San Francisco, built in 1994. The cities (San Francisco, Chicago, New York, and Washington, DC) that the NSF chose for the NAP locations currently hold four of the top sixth national ranks in bandwidth. This early NSF backbone infrastructure may have given these four cities a head start on bandwidth infrastructure that they have maintained since. There is a MAE in San Jose; MAE-San Jose. Several other private interconnection facilities are also clustered in the San Francisco Bay area (McIntee 2001).

The removal of the San Francisco-Oakland-San Jose cluster has a surprising effect upon the connectivity ranking. There are no west coast C/MSAs or any cities west of Chicago-Gary-Kenosha in the top 25 (Table 5-27). The network connectivity changes very little with the removal of this node, only 14%. The top of the ranking is comprised mainly of east coast cities and Florida cities. The hierarchy of links also exhibits an absence of the western half of the nation (Table 5-26). This is not surprising however, as the link raking tends to be dependent upon the node ranking. The pair removal scenario results show that the removal of this C/MSA and Dallas cause a 30% change in connectivity. Paired with Chicago-Gary-Kenosha causes even less change, 26% (Table 5-28). The most disruptive of the San Francisco-Oakland-San Jose pair scenarios was Atlanta, 70%.

Dallas-Fort Worth

Dallas-Fort Worth, though never at the top of Internet hierarchies, is a consistently strong C/MSA amongst the top ten Since 1997 Dallas has ranked within the top 10 cities in bandwidth (Kellerman 2002, p. 145). Currently (2003) Dallas-Fort Worth ranks

fourth in terms of bandwidth Dallas is also a top ten contender in fiber lit buildings (GeoTel 2003), colocation facilities (McIntee 2001), and wireless structures (Gorman & McIntee 2003). Dallas-Fort Worth also houses one of the original MAEs. Since the MAE was built in 1993 Dallas has been an important hub of Internet activity. Kellerman (2002, pp. 143-144) notes the number of direct Internet backbone connections between Seattle and Dallas. He discusses the flow of information (through the Internet backbone) is similar to the flows of people and commodities, and the connection is reflected between two high tech centers.

Table 5-27. U.S. city rankings based on the removal of San Francisco-Oakland-San Jose from the Internet backbone network

Rank	C/MSA	WRCI
1	Washington-Baltimore	6851.330
2	New York-Northern New Jersey-Long Island	5923.242
3	Philadelphia-Wilmington-Atlantic City	5622.690
4	Atlanta	5271.576
5	Jacksonville	4683.382
6	Orlando	4229.762
7	Tallahassee	3217.701
8	Tampa-St.Petersburg-Clearwater	3132.384
9	Miami-Fort Lauderdale	2940.492
10	Boston-Worcester-Lawrence	2598.633
11	Richmond-Petersburg	2375.295
12	Albany-Schenectady-Troy	1827.331
13	Pittsburgh	1772.286
14	Dallas-Fort Worth	1754.344
15	Houston-Galveston-Brazoria	1738.223
16	Cleveland-Akron	1662.245
17	Chicago-Gary-Kenosha	1477.704
18	Raliegh-Durham-Chapel Hill	1410.861
19	Charlotte-Gastonia-Rock Hill	1370.059
20	Hartford	1149.482
21	New Brunswick	1093.461
22	Harrisburg-Lebanon-Carlisle	967.556
23	Daytona Beach	918.508
24	Charlottesville	896.104
25	Providence-Fall River-Warwick	883.323

Note: NCI 92284.869; -13.5 % change; diameter 11; disconnects Honolulu

Upon the removal of Dallas-Fort Worth the network experiences only a -18% change in NCI. Los Angeles-Riverside-Orange County ranked 21st in the original network, but was pushed back to 25th when Dallas-Fort Worth was removed from the

network. Southern cities are again apparent in the top ten nodes (Table 5-30) as well as in the top ten links (Table 5-31). Table 5-32 shows the pair scenarios involving Dallas-Fort Worth. The scenario creating the most impact was Dallas-Fort Worth and Washington-Baltimore (70% NCI change). The scenario with the least impact is the joint removal of Dallas-Fort Worth and San Francisco-Oakland-San Jose, -30% (Table 5-32).

Table 5-28. Ranking of links based on the removal of San Francisco-Oakland-San Jose from the U.S. Internet backbone network

Rank	C/MSA	C/MSA	WRIC
1	Washington-Baltimore	New York-Northern New Jersey-Long Island	505.78
2	Washington-Baltimore	Philadelphia-Wilmington-Atlantic City	496.62
3	New York-Northern New Jersey-Long Island	Philadelphia-Wilmington-Atlantic City	434.71
4	Atlanta	Washington-Baltimore	377.66
5	Orlando	Jacksonville	318.01
6	New York-Northern New Jersey-Long Island	Atlanta	312.84
7	Jacksonville	Atlanta	308.69
8	Philadelphia-Wilmington-Atlantic City	Atlanta	289.00
9	Orlando	Atlanta	279.20
10	Washington-Baltimore	Jacksonville	274.78
11	Washington-Baltimore	Orlando	240.81
12	Tampa-St.Petersburg-Clearwater	Jacksonville	237.56

Table 5-29. Summary of node removal pairs with San Francisco-Oakland-San Jose from the U.S. Internet backbone network

	Network Connectivity Index (NCI)	% Change	Diameter	Disconnects
Initial Network Values	106726.1	N/A	11	N/A
Double node removal scenarios				
C/MSAs Removed with San Francisco-Oakland-San Jose:				
Dallas-Fort Worth	74793.029	-29.9	11	4
Atlanta	32276.762	-69.8	11	3
Chicago-Gary-Kenosha	78478.484	-26.5	11	7
New York-Northern New Jersey-Long Island	34622.124	-67.4	11	11
Washington-Baltimore	26449.206	-75.2	11	3

Chicago-Gary-Kenosha causes only a 16% change in connectivity with its removal (Table 5-7). Chicago-Gary-Kenosha ranks sixth in bandwidth (2003) but only 16th in the connectivity hierarchy of the U.S. Internet backbone network. This is surprising given that downtown Chicago houses the Chicago NAP, presumable the heart of the C/MSA's

Internet infrastructure and activity. The central location of Chicago-Gary-Kenosha is ideal for interconnection between the east and west coasts. Between 2000 and 2003 however Chicago-Gary-Kenosha dropped from second rank to sixth rank in bandwidth (Table 5-3).

Table 5-30. U.S. city rankings based on the removal of Dallas-Fort Worth from the Internet backbone network

Rank	C/MSA	WRCI
1	Washington-Baltimore	6758.665
2	New York-Northern New Jersey-Long Island	5955.750
3	Philadelphia-Wilmington-Atlantic City	5607.950
4	Atlanta	4824.337
5	Jacksonville	4221.449
6	Orlando	3729.340
7	Tallahassee	2903 088
8	Tampa-St.Petersburg-Clearwater	2718.558
9	Miami-Fort Lauderdale	2617.872
10	Boston-Worcester-Lawrence	2609 805
11	Richmond-Petersburg	2347 417
12	Albany-Schenectady-Troy	1845.394
13	Pittsburgh	1764.647
14	Cleveland-Akron	1668.388
15	Chicago-Gary-Kenosha	1448.552
16	San Francisco-Oakland-San Jose	1396.068
17	Raliegh-Durham-Chapel Hill	1359.701
18	Charlotte-Gastonia-Rock Hill	1285.194
19	Hartford	1160.276
20	New Brunswick	1099.274
21	Houston-Galveston-Brazoria	1008.614
22	Harrisburg-Lebanon-Carlisle	961.236
23	Providence-Fall River-Warwick	891.378
24	Charlottesville	887.285
25	Los Angeles-Riverside-Orange County	872.296

Note. NCI 87348.944; -18.1 % change, diameter 11; disconnects Amarillo

Chicago-Gary-Kenosha

Tables 5-33 and 5-34 show a node and link ranking with the usual top three cities leading the ranking; Washington-Baltimore, New York-Northern New Jersey-Long Island, and Philadelphia-Wilmington-Atlantic City. Atlanta fills the fourth rank in the hierarchy of node (Table 5-33). The ranks 5-10 are filled with Florida cities Daytona Beach also moves into the top 25, ranking 25 in the ranking of nodes. Table 5-35 shows the pair scenario with the least impact is Chicago-Gary-Kenosha and San Francisco-Oakland-San Jose. The NCI changes only -26%.

Table 5-31. Ranking of links based on the removal of Dallas-Fort Worth from the U.S. Internet backbone network

Rank	C/MSA	C/MSA	WRIC
1	Washington-Baltimore	New York-Northern New Jersey-Long Island	515.05
2	Washington-Baltimore	Philadelphia-Wilmington-Atlantic City	502.43
3	New York-Northern New Jersey-Long Island	Philadelphia-Wilmington-Atlantic City	444.97
4	Washington-Baltimore	Atlanta	360.68
5	New York-Northern New Jersey-Long Island	Atlanta	302.51
6	Orlando	Jacksonville	299.36
7	Atlanta	Jacksonville	286.09
8	Atlanta	Philadelphia-Wilmington-Atlantic City	277.27
9	Washington-Baltimore	Jacksonville	258.33
10	Atlanta	Orlando	255.49
11	Washington-Baltimore	Boston-Worcester-Lawrence	229.06
12	Jacksonville	Tampa-St.Petersburg-Clearwater	222.24

Table 5-32. Summary of node removal pairs with Dallas-Fort Worth from the U.S. Internet backbone network

	Network Connectivity Index (NCI)	% Change	Diameter	Disconnects
Initial Network Values	106726.1	N/A	11	N/A
Double node removal scenarios				
C/MSAs Removed with Dallas-Fort Worth:				
Atlanta	34543.589	-67.6	11	1
Chicago-Gary-Kenosha	73735.971	-30.9	11	5
San Francisco-Oakland-San Jose	74793.029	-29.9	11	4
New York-Northern New Jersey-Long Island	28376.720	-73.4	11	9
Washington-Baltimore	22294.329	-79.1	11	1

Link Removal

The majority of this research has focused on node removal scenarios. The importance of links to the network should not be ignored though the scope of this research concentrates on nodes. Link removal scenarios were performed as examples to show the impact of the removal of links, particularly the most critical links to the network's connectivity. The top-two links in the original network were removed in separate scenarios. The most important link in the network is the connection between New York-Northern New Jersey-Long Island and Washington-Baltimore. This link was

eliminated from the matrix, and then matrix multiplication was performed. The methodology is identical to that used the node-removal scenarios. Table 5-36 shows the resulting hierarchy and the drastically low WRCI values. With the removal of this critical link, the network experienced a -39% change in overall network connectivity. This change in connectivity is higher than the disturbance caused by five of the nine nodes removed (see Table 5-7). The diameter remained at eleven. The second link removal scenario eliminated the link between Philadelphia-Wilmington-Atlantic City and Washington-Baltimore. This removal scenario caused a -47% change in network connectivity. Table 5-37 shows the new hierarchy of nodes is lead by southern cities. Given these results, link removal should be further explored in future research.

Table 5-33. U.S. city ranking based on the removal of Chicago-Gary-Kenosha from the Internet backbone network

Rank	C/MSA	WRCI
1	Washington-Baltimore	6685.874
2	New York-Northern New Jersey-Long Island	5569.494
3	Philadelphia-Wilmington-Atlantic City	5419.064
4	Atlanta	5172.749
5	Jacksonville	4628.446
6	Orlando	4187.23
7	Tallahassee	3179.813
8	Tampa-St.Petersburg-Clearwater	3105.066
9	Miami-Fort Lauderdale	2908.806
10	Richmond-Petersburg	2325.675
11	Boston-Worcester-Lawrence	2223.290
12	Houston-Galveston-Brazoria	1721.301
13	Albany-Schenectady-Troy	1710.905
14	Pittsburgh	1679.970
15	Dallas-Fort Worth	1676.440
16	Cleveland-Akron	1420.086
17	Raliegh-Durham-Chapel Hill	1385.589
18	Charlotte-Gastonia-Rock Hill	1347.167
19	San Francisco-Oakland-San Jose	1333.891
20	Hartford	1064.678
21	New Brunswick	1042.676
22	Los Angeles-Riverside-Orange County	1003.476
23	Harrisburg-Lebanon-Carlisle	942.513
24	Daytona Beach	910.422
25	Charlottesville	877.552

Note: NCI 89811.660, -15.8 % change; diameter 11; disconnects Oakbrook, Rolling Meadows, Southfield, Glenview

Table 5-34 Ranking of links based on the removal of Chicago-Gary-Kenosha from the U.S. Internet backbone network

Rank	C/MSA	C/MSA	WRIC
1	Washington-Baltimore	Philadelphia-Wilmington-Atlantic City	490.61
2	Washington-Baltimore	New York-Northern New Jersey-Long Island	486.94
3	New York-Northern New Jersey-Long Island	Philadelphia-Wilmington-Atlantic City	415.13
4	Washington-Baltimore	Atlanta	376.17
5	Jacksonville	Orlando	317.70
6	Atlanta	Jacksonville	308.02
7	Atlanta	New York-Northern New Jersey-Long Island	301.69
8	Atlanta	Philadelphia-Wilmington-Atlantic City	284.53
9	Atlanta	Orlando	278.56
10	Washington-Baltimore	Jacksonville	274.43
11	Washington-Baltimore	Orlando	240.76
12	Jacksonville	Tampa-St.Petersburg-Clearwater	237.33

Table 5-35. Summary of node removal pairs with Chicago-Gary-Kenosha from the U.S Internet backbone network

	Network Connectivity Index (NCI)	% Change	Diameter	Disconnects
Initial Network Values	106726.1	N/A	11	N/A
Double node removal scenarios C/MSAs removed with Chicago-Gary-Kenosha				
Atlanta	31044.025	-70.9	11	4
Dallas-Fort Worth	73735.971	-30.9	11	5
San Francisco-Oakland-San Jose	78478.484	-26.5	11	7
New York-Northern New Jersey-Long Island	36665.223	-65.6	11	12
Washington-Baltimore	25592.043	-76.0	11	4

Conclusion

Chapter 5 has discussed the results of the weighted analysis. First, new measurements used for the weighted analysis were introduced. The WRCI was explained. A discussion of bandwidth in the U.S. followed. The past and present ranking of U.S. cities based on Internet backbone bandwidth were discussed. There is a direct, positive relationship between bandwidth ranking and the ranking of cities based on the WRCI. A positive relationship also exists between the number of redundant links to a C/MSA and the WRCI. The product of the U.S. Internet backbone network connectivity matrix was then discussed. This served as a comparison for each of the single node removal scenarios as well as the pair removal scenarios.

Table 5-36. Hierarchy of links based on the removal of the link between New York-Northern New Jersey-Long Island and Washington-Baltimore from the U.S. Internet backbone network

Rank	C/MSA	WRCI
1	Jacksonville	3857.969
2	Atlanta	3665.611
3	Orlando	3662.676
4	Washington-Baltimore	3139.499
5	Tampa-St.Petersburg-Clearwater	2780.899
6	Tallahassee	2644.031
7	Philadelphia-Wilmington-Atlantic City	2566.515
8	Miami-Fort Lauderdale	2558.483
9	New York-Northern New Jersey-Long Island	2300.831
10	Dallas-Fort Worth	1528.514
11	Houston-Galveston-Brazoria	1486.901
12	Boston-Worcester-Lawrence	1231.256
13	Richmond-Petersburg	1186.546
14	San Francisco-Oakland-San Jose	1126.970
15	Chicago-Gary-Kenosha	1054.832
16	Cleveland-Akron	963.949
17	Charlotte-Gastonia-Rock Hill	927.196
18	Pittsburgh	917.237
19	Los Angeles-Riverside-Orange County	879.797
20	Raliegh-Durham-Chapel Hill	846.332
21	Kansas City	838.742
22	Daytona Beach	809.868
23	Albany-Schenectady-Troy	786.086
24	Denver-Boulder-Greeley	739.475
25	Sacramento-Yolo	734.503

Note: NCI 65297 529; -65.6% change ; diameter 11; no disconnects

The results of the weighted analysis differed greatly from the unweighted analysis, as expected. Representing links using proper weights helped to provide a much more accurate hierarchy of connectivity for the links and nodes in the U.S. Internet backbone network. The hierarchy of the top 25 nodes in the unweighted analysis differed from weighted hierarchy (Table 5-38). Only about 50% of the nodes that ranked in the top 25 in the unweighted analysis also made the top 25 for the weighted connectivity analysis. The geographic distribution of the top 25 nodes in the unweighted network differed from that of the weighted (Figure 5-4). The unweighted top 25 nodes are more widely dispersed throughout the U.S. than the weighted nodes. All links are not equal, and

representing their weights based on actual bandwidth helps to determine what is actually happening in terms of connectivity.

Table 5-37. Hierarchy of links based on the removal of the link between Philadeiphia-Wilmington-Atlantic City and Washington-Baltimore from the U.S. Internet backbone network

Rank	C/MSA	WRCI
1	Jacksonville	3584.198
2	Orlando	3465.914
3	Atlanta	3228.757
4	Tampa-St.Petersburg-Clearwater	2646.970
5	Tallahassee	2455.980
6	Miami-Fort Lauderdale	2423.913
7	New York-Northern New Jersey-Long Island	2119.672
8	Washington-Baltimore	2009.109
9	Dallas-Fort Worth	1390.436
10	Houston-Galveston-Brazoria	1365.012
11	Boston-Worcester-Lawrence	1159.014
12	San Francisco-Oakland-San Jose	1014.290
13	Chicago-Gary-Kenosha	990.021
14	Philadelphia-Wilmington-Atlantic City	876.321
15	Richmond-Petersburg	810.093
16	Cleveland-Akron	809.550
17	Los Angeles-Riverside-Orange County	800.485
18	Kansas City	788.559
19	Charlotte-Gastonia-Rock Hill	783.297
20	Daytona Beach	770.546
21	Albany-Schenectady-Troy	747.656
22	Denver-Boulder-Greeley	697.089
23	Sacramento-Yolo	682.691
24	Raliegh-Durham-Chapel Hill	663.425
25	Fort Myers-Cape Coral	645.656

Note: NCI: 56341.534; % change -76.0; diameter 11; no disconnects

The importance of geography and the connectivity of the Internet backbone network was apparent based on the results of the weighted analysis. The node-removal scenarios revealed strong evidence of regional sub-networks within the U.S. Internet backbone network.

Atlanta emerged as a critical node for connecting Florida cities to the rest of the Internet backbone network. When Atlanta was removed from the network, all of the Florida cities in the network dropped significantly in terms of connectivity, none of them ranking within the top-ten. This is of little surprise, as Atlanta is centrally located and

ranks high in both connectivity and bandwidth. The removal of Atlanta caused a 60% drop in the NCI. The importance of Atlanta to the Florida peninsula illustrates the theory of sub-regional networks.

Table 5-38. Comparison of unweighted and weighted node hierarchies in the U.S. Internet backbone network

Rank	Unweighted Hierarchy	Weighted Hierarchy
1	Chicago-Gary-Kenosha	Washington-Baltimore
2	Washington-Baltimore	New York-Northern New Jersey-Long Island
3	Dallas-Fort Worth	Philadelphia-Wilmington-Atlantic City
4	San Francisco-Oakland-San Jose	Atlanta
5	Atlanta	Jacksonville
6	New York-Northern New Jersey-Long Island	Orlando
7	Kansas City	Tallahassee
8	Denver-Boulder-Greeley	Tampa-St.Petersburg-Clearwater
9	Los Angeles-Riverside-Orange County	Miami-Fort Lauderdale
10	St. Louis	Boston-Worcester-Lawrence
11	Sacramento-Yolo	Richmond-Petersburg
12	Seattle-Tacoma-Bremerton	Dallas-Fort Worth
13	Boston-Worcester-Lawrence	Albany-Schenectady-Troy
14	Houston-Galveston-Brazoria	Houston-Galveston-Brazoria
15	Cleveland-Akron	Pittsburgh
16	Miami-Fort Lauderdale	Chicago-Gary-Kenosha
17	Indianapolis	San Francisco-Oakland-San Jose
18	San Diego	Cleveland-Akron
19	Phoenix-Mesa	Raliegh-Durham-Chapel Hill
20	Pennsauken	Charlotte-Gastonia-Rock Hill
21	Tulsa	Los Angeles-Riverside-Orange County
22	Nashville	Hartford
23	Cincinnati-Hamilton	New Brunswick
24	Portland-Salem	Kansas City
25	Salt Lake City-Ogden	Sacramento-Yolo

Atlanta's importance as a regional hub was also illustrated in its paired-removal with Chicago-Gary-Kenosha, where the NCI drops 70.9%. Both cities are centrally located, though one in the north and one in the south and are dominant hubs in their regions. The removal of Atlanta paired with New York-Northern New Jersey-Long Island, Philadelphia-Wilmington-Atlantic City and Washington-Baltimore each caused a significant decrease in network connectivity as well, greater than 80%. Each of these cities serve as regional hubs, thus causing a double regional-hub impact with their removal.

Figure 5-4. Comparison of top 25 nodes based on URCI and WRCI values, with graduated symbols representing higher ranking cities.

One of the most surprising findings in the weighted analysis of the U.S. Internet backbone network is importance and prominence of Florida. Five Florida cities: Jacksonville, Orlando, Tallahassee, Tampa-St.Petersburg-Clearwater, and Miami-Fort Lauderdale ranked within the top-ten cities based on network connectivity. These cities have been experiencing rapid growth in telecommunication infrastructure, particularly in Internet backbone bandwidth Miami is the home of the newly-built NAP of the Americas which is considered by some to be the telecommunication gateway to South America. Another NAP is planned to be built in Jacksonville within the next year, further establishing the presence of Florida cities in the telecommunication hierarchy. Two Florida cities, Jacksonville and Orlando, were removed from the network to determine their removal impact. Each of these cities caused a significant drop in network connectivity with their removal, the NCI dropped 31% with the removal of Orlando, and 38% with the removal of Jacksonville. When both cities were removed simultaneous the network experienced a 43% drop in connectivity.

Further evidence of sub-regional networks is the significant drop in connectivity, -89% with the simultaneous removal of Jacksonville and Washington-Baltimore. This is the most significant drop in connectivity of all the node-removal scenarios performed in the weighted analysis. These are two predominant hubs in their defined regional sub-networks, causing a severe loss in connectivity with their removal due to the reduction of connection between regions.

The results of this analysis also implicate the importance of Southeastern, Northeastern, and Mid-Atlantic nodes to the overall network. Though some Western cities were included in the top-ranked nodes of the network, it can be generalized that they were not as critical to the overall connectivity to the network as were cities east of the Mississippi

The topic of sub-regional networks will be further explored and discussed in Chapter 6 using GIS as a tool. Chapter 6 will also explore various spatial and statistical relationships to help to explain the geographic distribution of the Internet backbone network. Based on these results, as well as the results of the weighted analysis, a list of the five most vulnerable links and nodes, in terms of network connectivity, will be compiled.

CHAPTER 6
UNDERSTANDING AND APPLYING THE RESULTS OF THE U.S. INTERNET BACKBONE NETWORK ANALYSIS

A regression analysis was employed to explain the variation of the relative connectivity of nodes and links in the U.S. Internet backbone network. This chapter also seeks to identify the location of the most important links and nodes based on a trend surface. In addition, this chapter presents a categorization system of nodes and links, based on the network concepts from graph theory and social network analysis for policy recommendation. The final sections in this chapter provide a summary and conclusions with recommendations for direction for future research.

Regression Analysis

This section develops a regression-based model as an attempt to explain interregional variations in the distribution of the Internet backbone network by C/MSA. Previous studies (Bebee & Gilling 1976, Sanders et al 1983) have shown definitive relationships between highly developed local and regional economies and highly developed infrastructure in the U.S. Although these studies are decades old, the theory can be successfully applied to Internet infrastructure. In addition, it is recognized that Internet infrastructure is also dependent upon the media industry, e-service industries, and financial industries.

Localized industrial profiles and population size was used by Dholakia and Bari (1994) to examine the relationship between economic development indices and telecommunications infrastructure Multiple regression analysis was employed with results that indicated a strong positive statistical relationship between investment in telecommunications infrastructure and economic development (Kotval 1999).

The research in this section will include a regression analysis, using graph-theoretic indices as a set of dependent variables and a selected set of independent variables. For the independent variables, several socio-economic characteristic statistics will be used including Gross National Product (GNP), bank deposits, percent of income from the manufacturing sector, percent of income from the service sector. Other independent variables include geographic description of the C/MSA's studied, for example, coastal vs, noncoastal locations, regional designators, and size. Demographic factors of the C/MSA's will also be used as independent variables including; population size and density, growth rates, income and per capita.

A stepwise regression model was employed as at least two of the variables used in the analysis displayed collinearity and would pose a problem in decreasing the precision of the individual estimated coefficients by interacting variance estimates. A database with descriptive statistics of 111 metropolitan areas was compiled (Table 6-1). The database included population, bank deposits, income, and the local economy's dependence on specific sectors, such as finance, insurance, and real estate (FIRE), as well as other factors that can be used as interactive variables in the model. Interactive variables were also used in a forward stepwise procedure, including interactive dummies (Table 6-1 for listing of variables). Many of the variables chosen, such as existing Internet infrastructure are known to influence the location of other types Internet infrastructure (Kotval 1999). Hypothetically, Internet backbone location decision would be influenced by the same factors as other types of infrastructure, such as colocation facilities, MAEs, and telephone switching equipment. Hence, it is hypothesized that infrastructure capital will be important in explaining variations in weighted network connectivity.

Table 6-1. Stepwise regression variable description table

Dependent variable	Description
Redundant links 2003	Number of redundant connections to an C/MSA 2003
Binary totals 2003	Total number of binary links 2003
Bandwidth 2003 totals	Amount of bandwidth connecting a C/MSA to Internet backbone network 2003
URCI 2003	Unweighted Relative Connectivity Index
WRCI 2003	Weighted Relative Connectivity Index
Bandwidth 2000	Amount of bandwidth connecting a C/MSA to Internet backbone network 2000
Fiber-lit buildings	Number of fiber-lit buildings within a C/MSA
Independent variable	**Description**
Redundant links 2003	Number of redundant connections to an C/MSA 2003
Binary totals 2003	Total number of binary links 2003
Bandwidth 2003 totals	Amount of bandwidth connecting a C/MSA to Internet backbone network 2003
URCI 2003	Unweighted Relative Connectivity Index
WRCI 2003	Weighted Relative Connectivity Index
Bandwidth 2000	Amount of bandwidth connecting a C/MSA to Internet backbone network 2000
Fiber-lit buildings	Number of fiber-lit buildings within a C/MSA
Colocation facilities	Number of Colocation Facilities within a C/MSA
MAE	Presence of a Metropolitan Area Exchange (MAE) within the C/MSA
NAP/IX point	Presence of a Network Access Point (NAP) or Internet Exchange (IX) Facility
Total # of cell towers	Number of Wireless Towers within a C/MSA
Auction value $	Value of FCC Wireless Auction for C/MSA
Rank-pop 2000	MSA ranking based on population size, Jan. 1999, taken from U.S. Census
Census pop 2000	MSA population based on the Census 2000 data
Pop change 1990-2000	Population change within an MSA from 1990-2000, Census 2000 data
bank deposits	Total (mil dol) bank deposits, 1996 Census data, SMDB
% inc.manufacturing	% Personal income manufacturing sector, 1994 Census data, SMDB
% inc services	% Personal income service sector, 1994 Census data, SMDB
% inc.retail trade	% Personal income retail trade sector, 1994 Census data, SMDB
% inc. fire fire	% Personal income finance, real estate, insurance 1994 Census data, SMDB
Government	% Personal income government sector, 1994 Census data, SMDB
Personal income per capita dollars	Personal Income per capita dollars,1994 Census Data, SMDB

Table 6-1. Continued.

Independent variable	Description
Personal income per capita rank	Personal Income per capita rank, 1994Census Data, SMDB
Constrained distance	Distance based accessibility based on the actual, existing connections
Distance row sum	Distance based accessibility between nodes, regardless of connection, unindexed, units in meters
Distance indexed	Distance based accessibility between nodes, regardless of connection, Indexed
Accessibility total	Sum of Row totals for each matrix until network reached diameter

* State and Metro Data Book (SMDB), United States Census 1997-1998

Silicon Valley colocation companies are experiencing a common demand from customers; they want their servers to stay near their headquarters (Templin 2001). Yet another location decision preference, this reiterates the idea that although firms could locate anywhere, they have location decision criteria, due in part to client demands. A modeling framework was employed to determine which variables might be the most influential in determining the geographic distribution of Internet backbone bandwidth and other types of telecommunication infrastructure in the U. S.

Table 6-2 shows the results of the Fiber-lit building model. The five most statistically significant variables in this model are number of colocation facilities within a C/MSA, presence of a MAE within a C/MSA, presence of a NAP/IX point with a C/MSA, population change from 1990-2000, and bank deposits within a C/MSA. With this model, these variables account for approximately 87% of the variation in the dependent variable (Table 6-3). The R-squared value in the overall model is relatively high, at 0.87, suggesting that this is an adequate model. All variables proved significant at the 95% confidence level, or higher The presence of a MAE proved to be significant. The colocation industry originated as an alternative to the congested MAEs and NAPs. A map of colocation facility distribution in the United States is shown in Figure 6-1. The distribution of MAEs, NAPs, and IX facilities are shown in Figure 6-2. The statistical results show that though a

Table 6-2. Forward stepwise regression results

Model and variables	Regression coefficient	Std. Error	t	Probability
Fiber-lit buildings				
Intercept	-18.095	18.254	-0.991	0.323
Colocation total	5.788	2.020	2.864	0.005
MAE	374.570	102.021	3.671	0.000
NAP/IX point	52.154	28.845	1.808	0.073
Pop change 1990-2000	170 479	91.379	1.865	0.064
Bank deposits	3.024	6.896	4.385	0.991
Redundant links 2003				
Intercept	5.237	2.009	2.606	0.010
Colo total	2 246	0.239	9.398	0.000
Binary totals				
Intercept	3.925	0.440	8.922	0.000
Colo total	0.300	0.020	15.011	0.000
Bandwidth 2003				
Intercept	41264.140	7762.670	5.315	0.000
Colo total	5985.740	352.908	16.961	0.000

Table 6-3 Significance of the overall models

Model	R square	Adj R sq	RMS error	F
Fiber-lit buildings	0.877	0.871	108.304	145.527
WRCI 2003	0.9388	0.9271	378.64	79.508
Redundant links 2003	0.788	0.784	18.055	198.958
Binary totals	0.676	0.673	3.989	225.341
Bandwidth 2003	0.727	0.724	70387.01	287.6815

Table 6-4. Spearman rank correlation results

	Dependent variables:		
Independent variables	Bandwidth 2003	URCI	WRCI
Bank deposits	0.7287**	0.7581**	0.6520**
% income manufacturing	0.0742	0.0319	0.0354
% income retail trade	- 0.4102**	-0.3758**	-0 3454**
% income fire	0 4412**	0.5227**	0.4770**
% income services	0.3896**	0.2942**	0.3868**
% income government	-0.2000*	-0.2080*	-0.1770*
Personal income per capita	0 4021**	0.4608**	0.3688**
Population change 1990-2000	0.0332	0.1353*	0.007
Census population 2000	0.7492**	0.8015**	0.6386**
Auction value	0.6809**	0.7221**	0.6323**

*Significant at 90% confidence level

**Significant at 95% confidence level

company may not be locating inside the physical MAE facility, they are still interested in reaping the benefit of existing infrastructure of a MAE. The location of a MAE within a metropolitan area greatly increases the number of colocation facilities that are housed in the metropolitan area. The concentration of colocation facilities proved to be significant at the 99% confidence level as well. This is expected, as the level of networks and bandwidth increase, so does the need for interconnection. Presence of a NAP/IX point within a C/MSA was significant at the 95% confidence level. Population change between 1990 and 2000 proved to be significant (at the 95% confidence level) indicating that population growth is a catalyst for the growth of fiber-lit buildings. Bank deposits proved significant and were positive related to the presence of a NAP or IX point

Financial hubs were the first places to attract the telecommunication industry after deregulation, and the significance of the change in population and heftier bank accounts confirms this (Finnie 1998). Metropolitan areas experiencing an increase in population that are dependent on the financial sector are seeing a greater number of colocation facilities. Growing populations in younger cities are demanding a growth in Internet infrastructure with the growth and development of the city, concurrent with older cities which are still growing in population and updating existing Internet infrastructure, or implementing new infrastructure in older buildings and neighborhoods.

Spearman rank correlation coeffecients were used to summarize the strength of the relationships between the Bandwidth 2003 totals, URCI 2003 values, and the WRCI 2003 values (Table 6-4) and the independent variables. The Spearman rank correlation test was used instead of the Pearson correlation test, as pretest analysis of the distributions indicated that they were "nonnormal." The results showed strong positive rank correlations between bank deposits and Bandwidth 2003, URCI 2003 values, and the WRCI 2003 values, all significant at the 95% confidence level. This also proved true for the percentage income derived from the retail trade sector, though all indices showed

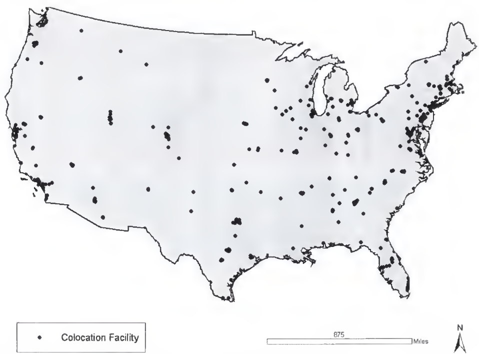


Figure 6-1. Distribution of colocation facilities in the United States (McIntee 2001)

Figure 6-2. Distribution of MAEs, NAPs, and Internet exchange points in the United States

strong negative correlations and were significant at the 95% confidence level. As retail trade increased, the Internet infrastructure indices decreased. No significant correlations existed between percentage income derived from the manufacturing sector and the indices, or between population change between 1990 and 2000 (Table 6-4). This implies that the Internet backbone network is not being driven simply by population growth, nor does it tend to gravitate towards those C/MSAs with economies driven by manufacturing.

Though population growth was not highly correlated with connectivity, Census population 2000 proved to have strong positive correlations with the indices. This proves that though growth may not be statistically associated with the indices, city size is important The percentage income derived from FIRE sectors showed moderately strong positive rank correlations significant at the 95% confidence level. The services sector was also positively correlated with each of the indices and significant at the 95% confidence level, but were a bit weaker than the correlation of other variables (Table 6-4). The government sector produced weak, negative correlations with the indices that were significant at the 90% confidence level. Personal income per capita showed moderately strong positive correlations with the indices. This represents the impact of market forces, suggesting that wealthier populations demanding more Internet infrastructure. Wireless auction values also showed strong positive correlations, all significant at the 95% confidence level (Table 6-4). This is an indication that the C/MSA distribution of wireless infrastructure replicating the pattern of geographic distribution of fixed telecommunication infrastructure. Figure 6-3 show s the distribution of wireless structures in the United States.

Kriging

Kriging is a fundamental tool in the field of geostatistics. Kriging is a method of interpolation named after a South African mining engineer, D. G. Krige, who developed the method in order to more accurately predict ore reserves. Interpolation is a method of increasing image resolution artificially. Kriging is based on the assumption that the

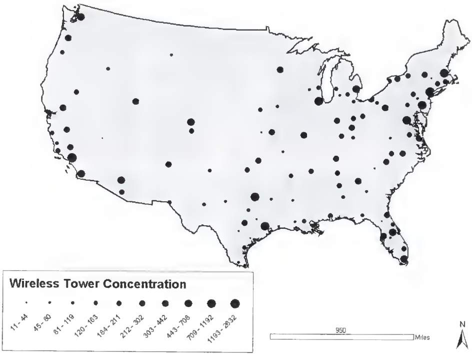


Figure 6-3. Concentration of wireless towers within C/MSAs in the United States

parameter being interpolated can be treated as a regionalized variable. A regionalized variable is intermediate between a truly random variable and a completely deterministic variable in that it varies in a continuous manner from one location to the next and therefore points that are near each other have a certain degree of spatial correlation, but points that are widely separated are statistically independent (Davis, 1986). Kriging is a set of linear regression routines that minimize estimation variance from a predefined covariance model. Kriging techniques with varying degrees of sophistication include ordinary kriging, universal kriging and indicator kriging.

The first step in ordinary kriging is to construct a variogram from the scatter point set to be interpolated. A variogram consists of two parts: an experimental variogram and a model variogram. The experimental variogram is determined by calculating the variance (g) of each point in the set in relation to each of the other points and plotting the variances versus distance (h) between the points. The variance is most commonly calculated as one half the difference in f squared.

Once the experimental variogram is computed, the next step is to define a model variogram, which is a simple function that models the trend in points. The variogram form shows that at small separation distances, the variance in f is small; points with similar f values are in close proximity to one another. After a certain level of separation, the variance in the f values becomes somewhat random and the model variogram flattens out to a value corresponding to the average variance.

The weights used in kriging are derived from the model variogram. The basic equation used in ordinary kriging is $F(x,y)=\sum_{i=1}^{n} w_i f_i$, where n is the number of scatter points in the set, fi are the values of the scatter points, and wi are weights assigned to each scatter point. This equation is essentially the same as the equation used for inverse distance weighted interpolation except that rather than using weights based on an arbitrary function of distance, the weights used in kriging are based on the model variogram.

By using the variogram in this fashion to compute the weights, the expected estimation error is minimized in a least squares sense. However, minimizing the expected error in a least squared sense is not always the most important criteria and in some cases, other interpolation schemes give more appropriate results (Philip & Watson 1986).

Surface Results

The results of the surface model are consistent with the research results in the unweighted and unweighted analyses. Figure 6-4 is the predicted surface based on the WRCI values. The J-Bessel model best fit the data and was employed for krigging the data. The surface is based on the 218 cities that were used in the U.S. Internet backbone network analysis. The values have been sorted into 20 classes. Figure 6-4 models the WRCI ranking based on dark shading over those areas that boast higher values, indicating an important node and most likely high bandwidth. These darker shaded areas are also representative of regional networks within the U.S. Internet backbone network. Florida and the Bo-Wash corridor are prominent, as well as Atlanta, Dallas, and the San Francisco Bay area. It should be noted that krigging is a process that is best applied to discrete data. The WRCI data is not continuous but was used to create this predicted surface.

The surface in Figure 6-4 was also used to help determine regional hubs in the network. Other factors that were used to define regional hubs include WRCI, geographical location to other nodes in the network, network connectivity removal impact, and bandwidth. Figure 6-5 shows the 37 nodes in the U.S. Internet backbone network that are defined as regional hubs. Figure 6-5 shows 60-mile buffers around each regional hub. The following section discusses an expanded regression analysis that incorporates additional variables that describe the geographic location and function of the network nodes. The next section also introduces interactive variables to the stepwise-regression analysis.

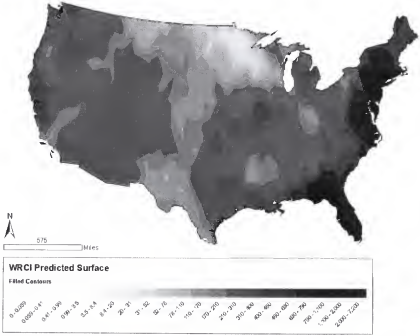


Figure 6-4. Predicted surface of the Weighted Relative Connectivity Index using C/MSA boundaries based on the WRCI values of 218 points, 20 classes employing the J-Bessel Model

Expanded Regression Analysis

A forward-stepwise regression procedure was run on the variables and dummies (including interaction terms) to identify those explanatory variables that explained the greatest variation in the dependent variable (WRCI). The inclusion/exclusion criteria was set at (.10 / .11, respectively), allowing for variables to enter (and be retained by) the model if estimated coefficients tested to be significantly different from zero at the 90% confidence level. Note that all coefficients in the above model tested significant at the 95% confidence level or higher

The model is not meant for prediction as much as it is to uncover statistical associations between the dependent variable WRCI and various Independent Variables and Dummies (to account for the variation in WRCI). Overall, the model is significant at the 99% confidence level, with an F-statistic of 79.5 (p=0.00000) (Table 6-5).

Table 6-5 Forward stepwise regression results

	Unstandardized Coefficients		Standardized Coefficients	
Model & Variables	Regression Coefficient	Std. Error	t	Probability
WRCI 2003				
Intercept (constant)	751.6638	338.2982	2.2219	0.028855
Redundant_Links_2003	41.7923	5.0252	8.3164	0.000000
Binary_Links_2003	-76.8782	24.6743	-3.1157	0.002478
MAE	-1618.515	308.4299	-5.2476	0.000001
Per_Inc_per_capita_99	-3.3306E-02	1.5235E-02	-2.1861	0.031463
Regional Hub	300.2518	113.6265	2.6424	0.009742
Coastal Location	-208.9440	98.4375	-2.1226	0.036596
SE	-1626.667	292.9650	-5.5524	0.000000
Population_change SE	2491.453	1065.780	2.3377	0.021672
Population_change Mid S	-1206.893	463.666	-2.6029	0.010846
X_Inc_Fire_SE	226.0468	36.3684	6.2154	0.000000
Pop_ Mid N	-9.7253E-05	4.4513E-05	-2.1848	0.031562
X_Inc_Fire_NE	80.0894	15.4020	5.1999	0.000001
NAPIX_Regional Hub SW	-1021.661	252.3771	-4.0482	0.000111
NAPIX_Regional Hub NE _Bank Deposit	3.7690E-02	5.3636E-03	7.0270	0.000000
Bandwidth per capita	1609 631	707.6377	2.2747	0.025357
Bandwidth per capita SE	4004.5740	993.1133	4.0323	0.000117

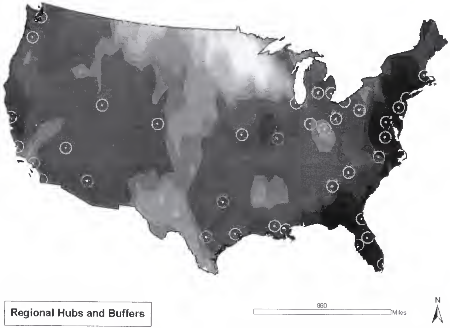


Figure 6-5. The 37 regional hubs in the U.S. Internet backbone network overlayed on a predicted surface using the WRCI values of 218 points.

Technical Notes. The correlation between WRCI and Network Accessibility is .99993. WRCI was chosen as the dependent variable as it represented a measure of a node's relative connectivity to all other nodes in the network at a distance of 10 - the network's known diameter. A sample of n=111 nodes was originally used. That sample was later reduced to 106 after the elimination of Greensboro, Kalamazoo, Lexington-Fayette, Pittsfield, and Roanoke due to lack of sufficient data for estimation.

Overall, the regression model did well to explain the variation in weight relative connectivity (WRCI), with an adjusted R-square of approximately 92.7% (Table 6-5). Several interesting results emerge from this analysis, revealing the importance of location and nodal function (from an economic and network standpoint).

The greater the redundancy of nodal connections, the greater the WRCI value as suggested by the positive coefficient associated with Redundant_Links_2003. This indicates that nodes with high levels of overall network connectivity (considering both direct and indirect links up to the diameter of the network) tend to have many redundancies in their direct binary connections. This suggests that market forces may be at work to reinforce the importance of critical nodes in the network allowing for the replication of critical links and multiple linkages in response to regional demand and national demand for bandwidth.

Note that there is a negative association between Binary_Totals_2003 (i.e., the number of direct connections) and WRCI. This suggests that a high level of local or regional connectivity do not necessarily translate into a high degree of network connectivity once all indirect links are considered. In other words, nodes of local prominence do not necessarily ensure their prominence when viewed from the perspective of the overall network. The negative coefficient for the variable MAE indicates that this type of infrastructure tends to be associated with nodes that are relatively less connected in the overall network. This suggests that the presence of a MAE does not necessarily

lead to a node having higher network connectivity. This is consistent with the WRCI ranking of cities that house MAE's, with the exception of Washington D.C. and New York. Figure 6-2 illustrates the distribution of MAEs, NAPs, and IX points.

Paris, Frankfurt, and a handful of U.S. cities are the only locations in the world that house MAEs. They are owned by MCI WorldCom, and originally operated using switching equipment that was not offered by many other peering points, which made them a unique type of data transfer and Internet backbone interconnection facility . For this reason, as well as commercial purposes, MCI named it's facilities Metropolitan Area Exchanges (MAE) and trademarked the name. Today, the original MAE locations are still operating, but using several types of switching equipment, similar to the other major interconnection hubs (NAPs and IX points) of the Internet. The CMSAs in the U.S. that currently house MAEs are Washington-Baltimore, San Francisco-Oakland-San Jose, Los Angeles-Riverside-Orange County, Dallas-Fort Worth, New York-Northern New Jersey-Long Island and Chicago-Gary-Kenosha. While Washington-Baltimore and New York-Northern New Jersey-Long Island boasted the highest WRCI value in the network, the other MAE locations ranked much lower. Dallas-Fort Worth ranked 12th in connectivity, Chicago 16th, San Francisco-Oakland-San Jose 17th, and Los Angeles-Riverside-Orange County 21st.

In general, coastal locations tend to have lower WRCI values once other considerations are taken into account. This implies that peripheral locations tend to have lower WRCI values all other things being equal. Bandwidth per capita (as a gross measure of local/regional transmission potential) tends to be positively related to WRCI (Table 6-5). This positive relationship is visibly stronger in the Southeast (see larger coefficient value for Bandwidth per capita SE in comparison to Bandwidth per capita). In terms of the Finance, Insurance, and Real Estate (FIRE) sector, a regional distinction emerges between northeastern and southeastern cities. While this sector's influence is shown to be positively related to WRCI in both regions, the presence of a strong FIRE

sector tends to promote higher WRCI values in the Southeast than in the Northeast. The bandwidth growth in the Southeast has been more significant than in any other region in the U.S. (Table 5-3). Several Florida cities including Orlando, Jacksonville, Tallahassee, Miami-Fort Lauderdale, Tampa-St.Petersburg-Clearwater have rapidly climbed the ranks in terms of Internet activity and infrastructure. This growth is also consistent with the population growth Florida is experiencing, the third fastest growing state in the nation (U.S. Census 2003). Other Southeastern cities, namely Raliegh-Durham-Chapel Hil, Charlotte-Gastonia-Rock Hill, and Atlanta, have experienced booms in Internet infrastructure, particularly Internet backbone bandwidth. Nonetheless, this result highlights the importance of the financial sector in explaining variations in WRCI in the most highly connected regions of the United States. It was surprising to see that the size of the government sector or population of the nodes in question did not enter into the model.

Note that the presence of NAPs in regional hubs of the SW tend to be associated with lower WRCI values in comparison to cities across the rest of the country. While regional hubs with NAPs in the northeast with larger bank deposits tend to have higher WRCI values. This again, reinforces the importance of the northeastern financial hubs in the explanation of regional differences in network connectivity. Wealthier populations place a higher demand on Internet infrastructure.

Nodes that serve as "regional hubs" tend to have a higher WRCI than nonhubs. This result is undoubtedly linked to their function as both collector and disseminator points in the network. Regional hubs serve as important information transfer junctions within a regional sub-network, and are important in terms of both relative location and function. A map of the regional hubs in the U.S. Internet backbone network is shown in the previous section (Figure 6-5).

Nodes located in the Southeast (SE) are typically associated with lower WRCI values, although high population growth centers in the SE (as represented by the variable

POP_change SE) tend to have relatively higher WRCI values. This result reinforces the notion that market forces may be at work to increase the relatively connectivity of nodes within geographic regions containing high population growth centers (such as would be the case for the corridor of Sunbelt cities that stretches from Atlanta to south Florida).

In contrast, high population growth centers in the southern Midwest tend to have lower WRCI values, indicating that the strength of the statistical association between population growth rates and WRCI weakens and reverses itself as one moves west of the Southeastern United States. This is consistent with the slow diffusion of backbone links that occurs from the east to west coast (O'Kelly and Grubesic 2002). Dallas, ranked 12th in terms of WRCI, is the first node in the network that is located west of the Mississippi River (Table 5-5). Houston is the next node with a noneastern position in the network, ranked 14th (Table 5-5).

The positive association between Bandwidth per capita SE and WRCI may indicate that infrastructure and connectivity is speculative in nature – explained by the anticipation of the continued growth of cities in the southeast. Moreover, saturation of higher bandwidth per capita and high WRCI values in the SE may be indicators that the southeastern region is emerging as part and parcel of the predominant fiber-optic corridor which spans from the Northern Atlantic to Southern Atlantic states. It is a corridor that bridges the cluster of cities that run from Boston to Washington D.C. down through the Southeast and on to Miami via regional hubs like Atlanta and Jacksonville. This is not surprising given the "centrality" of the East Coast in the global communications network.

The Bo-Wash corridor and the Florida peninsula are important in the global Internet backbone network. New York, Washington D.C, and Miami are three of the top-ten cities in intercontinental backbones. California is also an important international Internet backbone player, with Los Angeles and San Francisco also top-ten intercontinental backbone cities. These global Internet backbone hubs house marine cable landings,

bridging the Internet backbone intercontinental gap. Figure 6-6 illustrates the distribution of marine cable landings in the U.S.

Figure 6-6. Distribution of marine cable landings in the United States and Puerto Rico

Categorizing Nodes and Links for Policy Recommendation and Contribution

This section will relate directly to public policy contribution. Two systems for organizing the network have been developed: a categorization system for determining the most important nodes, and a system that determines the most important connectivity components in the network. A detailed categorization system is discussed in the first part of this section. This categorization system considers several types of descriptive characteristics that were derived from social network analysis (e.g., betweeness, closeness, boundary spanners) as well as connectivity indices, and other attributes. Criteria for determining the most important links and nodes in a network will be described,

and applied to the U.S. Internet backbone network in the second part of this section. The categorization system looks at nodality, including location, function, and various ranking. The connectivity and network importance system concentrates more on the indices of a node, but also accounts of the categorization of a node

This research has concentrated mainly on nodes, and this section is also geared towards nodes, however links are also discussed. These two systems can be used in conjunction with one another in a policy analysis process for policy recommendation, or they may be used separately, depending on the application. It is recommended however, that both systems be used for the policy analysis. This contribution is important as it can be used as an aid to direct allocation of funds and resources directly to those network components that are most important to the overall network, as their removal might cause the greatest disturbance to the network

The development of a categorization system for nodes and links would be helpful for understanding the roles of particular network nodes and allocating attention to ensure network health and protection. The categorization of nodes will be discussed first. Chapter 2 introduced various terms and concepts used in social network analysis, which will be directly applied to the results of this research to categorize network nodes and links

The nodes of the U.S. Internet backbone network play different roles in the network, some defined by their geographical location, others by their historical roles, and still others by the economic forces driving the metropolitan areas in which they are located. Nodes may be structurally equivalent, or completely unequal. Nodes with similar connective or physical characteristics may have vastly different functions in the overall network depending on their location or size

A classification scheme has been developed to characterize nodes based on their attributes and functions. Nodality, which is the degree a node's dominance, will be determined by a number of properties including; the closeness of node through distance-

based accessibility and constrained distance-based accessibility, WRCI, number and magnitude of disconnects and NCI decrease upon removal, degree (the number of direct connection a node has), betweenness, closeness, boundary spanners, gateways, and regional hubs. All in all, thirteen attributes are used to characterize nodes:

1. Distance based accessibility
2. Constrained distance-based accessibility
3. WRCI value
4. Removal impact disconnects (number and size of cities disconnected
5. Removal impact nci decrease (percentage change)
6. Number of direct connections
7. Betweenness
8. Closeness
9. Boundary spanners
10. Regional gateway
11. International gateway
12. Regional hub
13. International hub

Regional hubs clearly play an important role as prominent nodes in a sub-network. A map of sub-regional hubs and sub-regional networks is shown in Figure 6-6. The sub-regional hubs and networks have been defined based on the WRCI values, but also the percentage change that occurred in the removal scenarios. For example, the removal of Atlanta caused a significant drop in NCI, but also caused significant change throughout the Florida peninsula There is strong indication that Atlanta serves as a regional hub to other nodes in the network.

Table 6-6 will demonstrate the categorization of five nodes in the U.S. Internet backbone system. The nodes were randomly selected to illustrate categorization of several different types of nodes with different attributes, rather than simply categorizing the top-ten nodes based on connectivity. The first six attributes have numeric values, but the other seven attributes will be assigned a value between zero and three. A zero indicates a node does not exhibit the attribute, 1-node slightly exhibits the attribute, 2-node moderately exhibits the attribute, and 3-node strongly exhibits the attribute. The five nodes randomly chosen for categorization and ranking are Chicago, Kansas City, Miami, San

Francisco and Seattle (Table 6-6). The first six attributes were indexed by dividing the values in each column by the highest value in that column. The values were then multiplied by three and rounded to the nearest whole number (ranging from 0-3) (Table 6-7). Each of the thirteen attributes are now in a common value-system for totaling the values in order to determine the most important nodes in the network. Based on the totals in Table 6-7, Chicago is the most important node in the network, with a value of 34, followed closely by San Francisco, 32. Seattle ranked third (23), Miami fourth (22), and Kansas ranked last (18). Given that each of the attributes in this categorization system are considered with equal weight, Chicago would be the most valuable node to the network.

This categorization system can also be altered to place more emphasis or weight on a particular characteristic of a node. For example, if it was determined that betweenness of a node and regional hubs were twice as important as the other attributes in the system, the weight of each these characteristics should be doubled. If betweeness and regional hub attributes were doubled in Table 6-6 Kansas City, would have a higher value to the network. The system for determining the most important connective components of a network is less detailed, but equally as important. This methodology has been developed to determine the most important connective nodes and links in the Internet backbone network based on nodal measurements that include WRCI, percentage change with removal, accessibility and redundant links. This will be based largely on the WRCI values produced during the weighted analysis, but also considers the accessibility of nodes and links and the amount of disturbance their removal causes the overall network.

The four attributes for determining the most important connective nodes in a network are the following:

- WRCI
- Percentage change in NCI with removal
- Distance based accessibility
- Redundant links

Table 6-6 Categorization of five nodes in the U.S. Internet backbone network. Step one

Attributes	Accessibility	Accessibility index	Constrained accessibility	Constrained index	WRCI value	WRCI value index	Disconnects	Disconnects index	NCI decrease index	NCI decrease index	Direct connections	Direct Index	Betweenness	Closenes	Boundary spanners	Regional gateway	International gateway	Regional hub	International hub
Chicago	138124331	0.4	26530	0.5	1932	0.7	4	1.0	15 8	1.0	63	1.0	3	3	3	3	2	3	3
Kansas City	151683978	0.5	12723	0.3	1204	0.4	0	0.0	N/D	N/D	18	0.3	3	2	2	3	0	3	0
Miami	151337872	0.5	13169	0.3	2943	1.0	N/D	N/D	N/D	N/D	14	0.2	1	2	2	2	3	2	3
San Francisco	294772801	0.9	50218	1.0	1855	0.6	3	0.8	13.5	0.9	26	0.4	2	2	2	3	3	3	3
Seattle	301619314	1.0	21006	0.4	794	0.3	1	0.3	N/D	N/D	13	0.2	1	1	2	3	3	3	3

Table 6-7 Categorization of five nodes in the U.S. Internet backbone network: Step two

Attributes	Distance based accessibility index	Constrained distance-based accessibility index	WRCI value index	Removal disconnects index	Removal impact-NCI decrease indexed	Number of direct connections indexed	Betweenness	Closeness	Boundary spanners	Regional gateway	International gateway	Regional hub	International hub	Totals
Chicago	0.4 =1	0.5 =2	0.7 =2	1 =3	1 =3	1.0 =3	3	3	3	3	2	3	3	34
Kansas City	0.5 =2	0.3 =1	0.4 =1	0 =0	N/D =N/D	0.3 =1	3	2	2	3	0	3	0	18
Miami	0.5 =2	0.3 =1	1 =3	N/D =N/D	N/D =N/D	0.2 =1	1	2	2	2	3	2	3	22
San Francisco	0 9 =3	1.0 =3	0.6=2	0.8 =2	0.9 =3	0.4 =1	2	2	2	3	3	3	3	32
Seattle	1.0 =3	0.4 =1	0.3 =1	0.3 =1	N/D =N/D	0.2 =1	1	1	2	3	3	3	3	23

These attributes are then ranked (Table 6-8), and the ranks are summed to determine the most important nodes based on connectivity. Bandwidth is an important connective attribute, but is used to obtain the WRCI index, so it is not considered here as a separate attribute. The top-five nodes in the U.S. Internet backbone network based on the WRCI values have been categorized in Table 6-8. Washington-Baltimore ranks first in the connective-importance system, the Capital also ranked first in the ranking of nodes based on WRCI values. The second ranking node is New York-Northern New Jersey-Long Island, which also ranked second in the WRCI nodal ranking. The third rank however, is a tie between Philadelphia-Wilmington-Atlantic City and Atlanta. Philadelphia-Wilmington-Atlantic City had ranked higher than Atlanta in terms of the WRCI value, but Atlanta's removal disturbance was higher. The fifth rank was filled by Jacksonville, which had also ranked fifth in terms of WRCI values. The tie between Philadelphia-Wilmington-Atlantic City and Atlanta is indication that though the WRCI values are a good measurement of connectivity, the incorporation of other connective indices could show slightly different results.

Table 6-9 illustrates categorization of links based on connective properties. The two links that were used in the link-removal scenarios, New York-Northern New Jersey-Long Island to Washington-Baltimore and Philadelphia-Wilmington-Atlantic City to Washington-Baltimore were used as examples (Figure 6-7). The attributes used to determine the most important links based on connectivity were slightly different than those used for the node categorization. The link attributes considered are the following:

- WRCI
- Distance
- Number of carriers comprising link (indicating redundancy)
- Removal impact (% decrease in NCI)

The WRCI value illustrates the connective property of a link. Distance is important because the length of a link is related to it's vulnerability The distance between New York and Washington D.C. is roughly 266 miles long, which translates to 266 possible miles of

Table 6-8. Criteria for determining the most important connectivity nodes in a network for resources allocation

WRCI Ranking (metropolitan area and WRCI value)		Rank based on WRCI	% Change in NCI upon removal	Ranked based on NCI change	Distance based accessibility	Ranked based on accessibility	Redundant links	Rank based on redundant links	Ranking yotals	Final ranking for policy recommendation
Washington-Baltimore	7178.152	1	-69.8	1	23162.00	1	245	1	4	1
New York-Northern New Jersey-Long Island	6294.686	2	-61.3	2	3637.50	3	183	2	9	2
Philadelphia-Wilmington-Atlantic City	5928.210	3	-53.8	4	2084.00	4	50	4	15	3
Atlanta	5444.191	4	-59.1	3	19051.00	5	105	3	15	3
Jacksonville	4650.768	5	-38.5	5	4893.00	2	40	5	17	5

Table 6-9. Criteria for determining the most important links in a network for resources allocation

Link	WRCI	WRCI ranking	Distance	Ranking based on distance	Number of carriers comprising link	Ranking based on number of carriers	% change in NCI upon removal	Ranking based on NCI	Ranking Totals	Final ranking for policy recommendation
New York-Northern New Jersey-Long Island to Washington-Baltimore	266.775	1	204 miles	1	17	1	-65.6	2	5	1
Philadelphia-Wilmington-Atlantic City to Washington-Baltimore	242.090	2	126 miles	2	16	2	-76.0	1	6	2

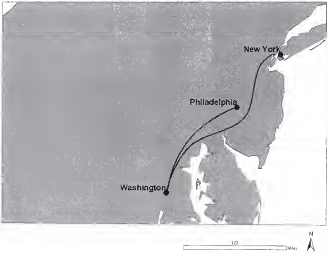


Figure 6-7. The most important links in the U.S. Internet backbone network based on WRCI

susceptibility for the connection between these two nodes. The number of carriers that comprise a link indicate the amount of redundancy a links has. For example, if Sprint, MCI WorldCom, and Verizon each had separate connections between New York and Washington D C., the redundancy would be three. Given that the physical connections had three different routes or paths the redundancy would help to ensure the health of the connection between these two cities. However, if these three connections were bundled in the same pipe, the redundancy would have significantly less importance, as a disturbance to this pipe would likely affect all three carriers simultaneously. The impact of a link's removal is also considered, by using the percentage decrease in NCI after a node has been eliminated from the network.

The ranking totals in indicate that the link between New York-Northern New Jersey-Long Island and Washington-Baltimore is a more important connective property to the network than the link between Philadelphia-Wilmington-Atlantic City and Washington-

Baltimore. Thus, it is recommended that resource allocation give priority to the connection between New York-Northern New Jersey-Long Island and Washington-Baltimore.

Important Nodes in the Internet Backbone Network

Based on the results of this network analysis, as well as the results of previous researchers who have studied various types of Internet infrastructure, the ten most important nodes to the U.S Internet backbone network are San Francisco, Kansas City, Chicago, Miami, Jacksonville, Atlanta, Dallas, Washington D.C., New York, and Philadelphia (Figure 6-8). This selection of nodes is based solely on the domestic function and importance of these nodes

San Francisco, though lagging in WRCI, is an important hub in California in terms of other Internet backbone measurements. San Francisco has been a consistent leader in terms of bandwidth and redundant links (Table 5-1, Table 5-3). This city is one of few that houses a MAE, and also leads in terms of colocation facilities (Figure 6-1, Figure 6-2).

Kansas City has been chosen based on it's central location to the rest of the nodes in the network, and has also grown significantly in bandwidth. Chicago, another mid-western node, is a top-ranked city in terms of bandwidth, as well as Internet backbone network interconnection activity. Chicago has a high concentration of colocation facilities, a MAE, NAP, and IX facilities. Historically, Chicago has been important hub of Internet activity. A MAE is also located in Dallas. The network experienced an 18% decrease in NCI with the removal of Dallas, a regional hub to Texas nodes (Table 5-8).

Miami and Jacksonville have both experienced significant growth in Internet bandwidth (Table 5-3) Jacksonville caused a surprisingly high disruption to the NCI upon it's removal (38% decrease) and is ranked high in WRCI. Miami has become an important hub in the past five years, particularly when the Miami NAP was built. This research has shown the new importance of Florida to the Internet backbone network, with emphasis on these two cities.

Figure 6-8. The most important nodes in the U.S. Internet backbone network considering domestic roles

Atlanta serves as an important hub to the network, and was ranked fifth in bandwidth, and fourth in connectivity (Table 5-3, Table 5-5). Atlanta's removal caused a major drop in network connectivity (Table 5-8).

Three nodes in the Bo-Wash megalopolis were considered ten of the most important nodes to the network: Washington D.C , New York, and Philadelphia. Washington D.C. has a high concentration of bandwidth, colocation facilities, a MAE, NAPs, and IX facilities. The removal of the Capital caused more disruption than any node removal scenario, the NCI decreased nearly 70% (Table 5-8). New York's removal caused a 61% decrease in NCI, which is partially explained by the fact this city ranks second in both bandwidth and WRCI. Philadephia was a surprising addition to the top five cities based on WRCI, due to it's lower ranking of bandwidth (13th).

Figure 6-9 illustrates the ten nodes considered most important to the Internet backbone network, considering both domestic and international function and importance. New York, Philadelphia, Washington D.C., Atlanta, Jacksonville, Miami, Dallas, and San Francisco were again chosen as ten of the most important nodes to the network. However, Kansas City and Chicago were replaced by Seattle and Charlotte. The coastal location of Seattle and it's prominence to the international Internet network Seattle houses marine cable-landings (Figure 6-6) the last stop for trans-continental data traveling to Asia. Charlotte's recent boom in industry and population, as well as Internet infrastructure has earned this city a place in the top ten most important nodes.

Summary and Conclusions

The intent of this research was to explore the U.S. Internet backbone network, specifically the connectivity of nodes and links. As has been discussed, the vulnerability of the nation's critical infrastructure is a pressing issue of homeland security. Though there have been several publications concerning the geographic distribution of the Internet backbone, this is the first research project that has used the 2003 Internet backbone data

Figure 6-9. The most important International nodes in the U.S. Internet backbone network considering domestic and international roles

set. This research is also the first to address the connectivity of the current Internet backbone using matrix multiplication. The importance of exploring the Internet backbone and the repercussions of disturbance of a node or link to the connectivity of the network has been described, and this dissertation sought to address the topology of the Internet backbone network for policy recommendation and contribution.

The main objective of this research was to determine which U.S. cities were the most important nodes in the network based on their connectivity. Matrix multiplication has helped to answer the research questions which were addressed in Chapter 4, Chapter 5, and Chapter 6. This research indicated the most critical nodes and links in the U S. Internet backbone network based on matrix multiplication and network analysis. Chapter 4 indicated the most connected nodes based on an unweighted analysis. Chapter 5 produced a more accurate ranking of connected cities by adding actual bandwidth values to the network. The five most connected nodes in the Internet backbone network based on the WRCI values are Washington-Baltimore, New York-Northern New Jersey-Long Island, Philadelphia-Wilmington-Atlantic City, Atlanta, and Jacksonville

The results of this research will aid policy-makers in the planning and protection of the U.S. Internet backbone network The research has determined which cities are the most important to the overall network, and which cause the largest amount of connectivity disruption when they are removed. Based on the WRCI index values obtained from the results of matrix multiplication, the connectivity of the nodes in the network were determined. Early research on the Internet backbone network and the exploration of spatial distribution of other types of Internet infrastructures indicated that these infrastructures tend to cluster in the same metropolitan areas This means that the same metropolitan areas would lead the ranking of various types of infrastructures; colocation facilities, telephone switches, NAPs, MAEs, IX points, POPs, fiber-optic bandwidth. The ranking similarity for different types of infrastructure has been proven to be true, based on

the comparison of results of various types of Internet infrastructure and activity (see Comparison C/MSA Internet Rankings, Table 4-5)

It was hypothesized that the metropolitan areas that ranked highest in terms of connectivity would also rank the highest in terms of interconnection facilities, specifically colocation facilities, but also the large interconnection hubs; IX points, NAPs, and MAEs. This hypothesis is supported. The metropolitan areas leading in WRCI, also house large numbers of interconnection facilities. The high, positive correlation between colocation facilities and the WRCI has been statistically confirmed using stepwise regression analysis. This further confirms that colocation facilities tend to cluster in those areas which house fiber

It was further hypothesized that the most highly connected nodes would also be those metropolitan areas which housed the most Internet bandwidth. The results of this dissertation do indicate that those metropolitan areas which rank high in Internet bandwidth also rank high in terms of network connectivity, based on the WRCI values (see Tables 5-3 & 5-5). Washington-Baltimore and New York-Northern New Jersey-Long Island rank first and second respectively in terms of both bandwidth and WRCI. Philadelphia-Wilmington-Atlantic City ranks third in terms of connectivity. This city ranks a little lower in bandwidth, but still remains strong, ranking 13th. Atlanta ranks fourth in connectivity and fifth in bandwidth. Jacksonville, the surprising newcomer to the top of the ranking is the fifth most connected city in the network Jacksonville has experienced huge growth in Internet bandwidth, moving from the 37th rank it held in 2000 to eighth by 2003.

The dissertation also intended to identify critical locations of Internet activity. It was theorized that locations heavy in Internet activity would be those places housing both the most Internet bandwidth and interconnection facilities The ranking of cities based on WRCI determined by this research is consistent with the city rankings of other types of

Internet infrastructures and activity (Table 4-5). Map 6- shows those cities defined in this research as 'hot-spots' of Internet activity.

Finally, it was hypothesized that there is great variability in the disruption effects of different links and nodes in the network based on the connective properties of the network. Both single-node and pairs of nodes removal scenarios were performed in Chapter 4 and Chapter 5. The results of the removal scenarios confirmed the hypothesis of disruption variability, particularly in the removal scenarios performed in Chapter 5. Chapter 5 showed that the removal of the top ranking nodes in the network significantly effected the Network Connectivity Index (NCI). The removal scenarios involving Atlanta and Jacksonville also demonstrated the large disruption a regional network experiences with the loss of a hub.

The link-removal-scenarios revealed surprising results. The removal of the top-ranked link in the network (between Washington-Baltimore and New York-Northern New Jersey-Long Island) caused significant disruption to the network, lowering the NCI by nearly 40%. The removal of the second-ranked link (between Philadelphia-Wilmington-Atlantic City and Washington-Baltimore) caused an even larger disruption, lowering the connectivity of the network roughly 50%. The link-removal scenarios caused nearly as much of a disturbance to the connectivity of the network as did the node-removal scenarios. The link removal scenarios demonstrated that they are as critical to the connectivity of the network as the network nodes.

This research confirms that telecommunications technologies are not a replacement for personal interactions but an enhancement. Moss (1998), Wheeler, Aoyama and Warf (2000), and Sassen (2000) have expressed that telecommunications are not a substitute for face-to-face interactions but that personal contact and telecommunications are complements, and they predict cities would become more important as information technology improves. The empirical evidence highlighted in this dissertation supports their opinions and predictions. The results also suggest the city is not on the decline;

sophisticated infrastructure is being built primarily within the largest, most affluent metropolitan areas in the United States. The growth in bandwidth between 2000 and 2003 (see Table 503) reaffirms the city's demand for Internet infrastructure.

Directions for Future Research

The ranking of nodes and links in the Internet backbone network has been determined. The weighted node-removal scenarios showed varied connectivity impact with the removal of different nodes from the network. The removal of the top-five nodes based on connectivity and the top-five nodes based on bandwidth caused drops in the Network Connectivity Index (NCI) ranging from 13.5%-69.8% (see Table 5-8). The removal of the top-two links in the network caused significant decrease in the NCI as well. This research has concentrated mainly on the removal of nodes from the network, but it is apparent from the few link-removal scenarios performed that the removal of a link from the network can be equally significant to the removal of a node, if not more so.

With the results of the link removal scenarios being as significant and substantial as they are, the first suggestion for future research is further exploration of the links of the Internet backbone network. Little research has been published on the effects of removing or disabling links in the U.S. Internet backbone network. The importance of links must be addressed in terms of a link's prominence to the network. Redundancy of a link is an important attribute to consider To completely sever or disconnect direct linkage between two cities, it is likely that multiple connections must be impaired For example, if twenty-five Internet backbone companies owned direct connection fiber links between New York and Chicago, in order to completely sever direct connection between these two cities, each of these separate links must be impaired. However, it is likely that the fiber-optic Internet bandwidth cable are located in close proximity to one another, if not in the same physical pipe. Many Internet backbone links follow Interstate highways and other major roads, as well as the railway system. For this reason, and many others, locating and disabling a fiber

link would prove to be a much easier task than removing a node from the network. Because the removal of links can potentially cause an equal amount of disruption to a network's connectivity as the removal of a node, and they are more feasible to damage or destroy, attention should be paid to further exploration of the most important links in the Internet backbone network.

Further research of links should include link removal impact upon the overall network, but attention must also be directed towards a link's importance to sub-regional networks. This research has defined several sub-regional networks within the U.S. Internet backbone network. The removal of a link that is critical to a subregional network could have significant impact to the connectivity of the sub-regional network to which it is connected. Sub-regional networks that are geographically adjacent will theoretically have impacts upon one another, especially if they are sharing a common link.

It could be argued that these nodes and cross-over links are the most vulnerable, leading to a conclusion that more attention should be turned toward adding redundancy and additional crossover links between regional sub-networks to prevent any one sub-network from being disconnected from the others in the network, not to mention more circuits for important peripheral players that link the domestic system to the global network. Another reason why the Central and North central nodes in Florida have gained prominence is that they are part and parcel of possible circuits that connect the Southeastern regional sub-network of the U.S. where Atlanta is the primary node to the rest of the domestic network. Hence, intermediate nodes like Orlando and Jacksonville become important as go-between nodes to regional, domestic, and global network, given their Florida connections.

It is recommended that link-removal scenarios be performed using a method similar to the matrix multiplication employed in this research. New network ranking can be obtained by simulating the network connectivity with the absence of a link, or pair of links.

The hierarchy of nodes and links, as well as the connectivity indices for the network, nodes, and links, would aid in further understanding the U.S. Internet backbone network for protection and planning to ensure a healthy network.

The U.S. Internet backbone should also be explored with international scope. Very little research has explored the Internet from a global view-point. Kellerman (2002) and Telegeography (2000, 2001) have determined the most important international links, nodes and estimated traffic statistics. However, the there is much to be investigated to better understand the Internet backbone within an international context. Nodes and links that are top-ranking domestically might not be as prominent to the international Internet backbone network. Network components might be very important global hub or connections, yet they are not as important to the national network. With the matter of limited resources and funds, should protection and insurance emphasis be placed on those links and nodes which are most important to the International Internet backbone network, or those that are critical to the national Internet backbone network.

REFERENCES

Abbate, J. (1999) Inventing the Internet. Cambridge, MA: MIT Press.

Abello, J., Pardalos, P.M., & Resende, M.G.C. (1999) External Memory Algorithms, in J. Abello and J. Vitter (Eds.), DIMACS Series in Discrete Mathematics Theoretical Computer Science, p. 119. Washington, DC: American Mathematical Society.

Abler, R. (1991) Hardware, Software, and Brainware: Mapping and Understanding Telecommunications Technologies, in S. D. Brunn & T. R. Leinbach (Eds.), Collapsing Time and Space, pp. 31-48. Hammersmit, UK: HarperCollins Academic.

Adamic, L. A. (1999) Proceedings of the Third European Conference, ECDL '99, p. 443. Berlin: Springer-Verlag.

Adamic, L. A., & Huberman, B. A. (1999) Growth Dynamics of the World Wide Web. Nature 401, 131.

Adamic, L. A., & Huberman, B. A. (2000) Power Law Distribution of the world Wide Web. Science, 287, 2115.

Adams, R. B. (June 1971) "Airline Connectivity Matrix: Growth and Disequilibrium," a paper presented at the annual meeting of the Mid-Continental Division of the Regional Science Association, Winnipeg, Manitoba.

Aiello, W., Chung, F., & Lu, L. (2000) Proceedings of the 32nd ACM Symposium on the Theory of Computing (ACM, New York), p. 171.

Albert, R., & Barabasi, A-L. (2000) Topology of Complex Networks. Local Events and Universality. Physical Review Letters 85, 5234.

Albert, R., & Barabasi, A-L. (2002a) Linked. Penguin Group, New York, New York.

Albert, R., & Barabasi, A-L. (2002b) Statistical mechanics of complex networks. Reviews of Modern Physics, 74, 47-97.

Albert, R., Jeong, H., & Barabasi A-L. (1999) The Diameter of the World Wide Web. Nature 401, 131.

Aoyama, Y. (2001) The Information Society, Japanese Style: Corner Stores as Hubs for E-Commerce Access, Leinbach & Brunn (Eds.), Worlds of E-Commerce (Chapter 6). Chichester West Sussex, England John Wiley & Sons.

Atkinson, R. & Gottlieb, P. (2001) The Metropolitan New Economy Index. www.neweconomyindex.org. Washington, DC: Progressive Policy Institute.

Barabasi, A-L., & Albert, R. (1999) Emergence of Scaling in Complex Networks. Science 286, 509.

Barabasi, A-L., Jeong, H., Ravasz, E., Neda, Z., Schubert, A., & Vicsek, T. (2001) Preprint con-mat/0104162.

Beniger, J. R. (1986) The Control Revolution: Technological and Economic Origins of the Information Society. Cambridge, MA: Harvard University Press.

Bernhardsen, T. (2002) Geographic Information Systems: An Introduction. John Wiley and Sons, New York, New York.

Berry, B. (1960) "An Inductive Approach to the Regionalization of Economic Development," University of Chicago, Department of Georgraphy Research Paper No. 62, Essays on Georgraphy and Economic Development, Ed. N. S. Gingsburg, Chicago: University of Chicago Press.

Bleyer in Business Florida (2004) http://www.businessflorida.com/advantages/06.asp .

Boardwatch Magazine (1999) The Internet—What Is It? Boardwatch Magazine's Directory of Internet Service Providers. Penton Media. http://boardwatch.internet.com/isp/summer99/

Boardwatch Magazine (2000) The Internet—What Is It? Boardwatch Magazine's Directory of Internet Service Providers, 12th edition. Cleveland: Penton Media.

Brunn, S. D., & Leinbach, T. R. (Eds.) (1991) Collapsing Space and Time: Geographic Aspects of Communication and Information. New York: Harper Collins Academic.

Button, K., & Taylor, S. (2001) Toward an Economics of the Internet and E-Commerce, in Leinbach & Brunn (Eds.), Worlds of E-Commerce (Chapter 2). West Sussex, England: John Wiley & Sons.

Cairncross, F. (1997) The Death of Distance, Boston: Harvard Business School Press.

Camacho, J., Guimera, R., & Amaral, L A N. (2002a) Robust Patterns in Food Web Structure, Phys. Rev. Lett 88, 228102.

Camacho, J., Guimera, R., & Amaral, L.A.N. (2002b) Analytical Solution of a Model for Complex Food Webs, Phys Rev E 65, 030901R.

Careless, J. (2003) Teleports Battle Through Blackout. Telecom Web: Via Satellite.com http://www.telecomweb.com/satellite/viasatellite/current/exclusive21.htm

Castells, M. (1989) The Informational Cit: Information Technology, Economic Restructuring and the Urban-Regional Process. Oxford: Basil Blackwell

Castels, M. (1993) The Informational Economy and the New International Division of Labor, in M. Carnoy, M. Castells, S. S. Cohen, & F. H. Cardoso (Eds.), The New Global Economy in the Information Age: Reflections on Our Changing World, pp. 15-43. University Park, PA: Pennsylvania State University Press.

Castells, M. (1996) The Rise of the Network Society. The Information Age, vol. 1. Oxford, Blackwell (Chapters 5,6,7,8,11,12,16)

Castells, M. (1997) The Information Age, vol. 2· The Power of Identify. Cambridge. Blackwell. (Chapter 12)

Castells, M. (1998) End of Millennium. The Information Age: Economy, Society and Culture, vol. 3. Oxford: Blackwell. (Chapter 11)

Castells, M. (1999) Grassrooting the space of flows, Urban Geography, 20 (4), May, pp. 294-302 (Chapter 6)

Centers for Disease Control (2003) Outbreak of Severe Acute Respiratory Syndrome Worldwide. http://www.iec.org/online/tutorials/opt_net/topic01.html March 21, 2003.

Clinton, W. (1998) www.unit.net/en/news/wto3.cfm. Bill Clinton Death of Distance address, 1998, May in Glattburg Web hosts: Unit.Net AG.

Cobb, S. (2001) Internet Economies and Online Recruiting Industry, in T. R. Leinbach & S. D. Brunn (Eds.), Worlds of E-Commerce (Chapter 7). West Sussex, England: John Wiley & Sons

Coe, N., & Yeung, H. W. (2001) Grounding Global Flows: Constructing an E-Commerce Hub in Singapore, in T. R. Leinbach & S. D. Brunn (Eds.), Worlds of E-Commerce (Chapter 8). West Sussex, England: John Wiley & Sons.

ColoSolutions (2001) www.colosolutions.com

Cukier, K. N. (1998) The Global Internet: A Primer, TeleGeography 1999. Washington, DC: Telegeography.

Dicken, P. (1994) Global-local Tensions: Firms and States in the Global Space-economy. Economic Geography, 70, pp. 101-28 (Chapter 11).

Dicken, P. (1998) Global Shift: Transforming the World Economy, 3rd edition, Chapters 8, 11. London: Paul Chapman.

Dodge, M., & Kitchin, R. (2001) Mapping Cybersapce. London: Routledge.

Economist (2001, August 9) Special Report: Geography and the Net. Putting it in its place. http://www.economist.com/printedition/displaystory.cfm?story_id=729808

Environmental Systems Research Institute (ESRI) (2002) Working with the Geodatabase· Powerful Multiuser Editing and Sophisticated Data Integrity. ESRI White Paper February, 2002. Redlands, California

Evans-Cowley, J., Malecki, E. J., & McIntee, A. (2002) Planning Responses to Telecom Hotels: What accounts for increased regulation of co-location facilities? Journal of Urban Technology, Volume 9: December.

Faloutsos, M., Faloutsos, P , & Faloutsos, C. (1999) Sampling Internet Topologies: How Small Can We Go. Computating Communication review 29, p. 251

Federal Communications Commission (2000) Deployment of Advanced Telecommunications Capability: Second Report. www.fcc.gov/bureaus/common_carrier/orders/2000/fcc00290 pdf

Ferrer i Cancho, R., & R V. Sole, R. V. (2001) The Small World of Human language. Santa Fe Institute working paper 01-03-016

Finnie, G. (1998) Wired Cities, Communications Week International. www.totaltele.com/cwl.

Forman, R., & Spearling, D. (Eds.) (2002) Road Ecology: Science and Solutions. Washington, DC: Island Press.

Gabel, D. (1996) Private Telecommunications Networks: An Historical Perspective, in E. Noam & A. Nishuilleabhain (Eds.), Private Networks Public Objectives, pp.35-49. Amsterdam: Elsevier

GAO February 2003 Potential Terrorist Attacks: Additional Actions Needed to Better Prepare Critical Financial Market Participants Report to Congressional Requesters http://www.gao gov/new.items/d03251.pdf - 5594.9KB - GAO Reports - Telecommunications: 4, September: 5, 11: 10, Telecommunications September 4

Garrison, W. L. (1956). Allocation of Road and Street Costs, Part IV. The Benefits of Rural Roads to Rural Property, State Council for Highway Research, Seattle, WA, June 1956.

Garrison, W. L. (1960): Connectivity of the interstate highway system. Regional Science Association, Papers and Proceedings, 6, pp.121-37. (3.5.1, 9.7.1).

Garrison, W. L., & Marble, D. F. (1961) The Structure of Transportation Networks Washington, DC: U.S. Department of Commerce, Office of Technical Services.

Geo-Tel (2002) GeoTel.Telecommunications Geography. http://www.geo-tel.com/

Gilder, G. (1995) Gilder Meets His Critics. Forbes ASAP, February 27, p. 56.

Gorman, S. P., & McIntee, A (2003) Tethered Connectivity: The Spatial Distribution of Wireless Technologies. Environment & Planning A 35 (7) pp.1157 - 1171.

Govindan, R., & Tangmunarunkit, H. 2000, in Proceedings of IEEE INFOCOM 2000, Tel Aviv, Israel (IEEE, Piscataway, N.J.), Vol. 3, p. 1371.

Graham, S. (1999) Global Grids of Glass: On Global Cities, Telecommunications and Planetary Urban Networks Urban Studies 36(5-6), pp.929-949.

Graham, S. (2000) Constructing Premium Network Spaces: Reflections on Infrastructure Networks and Contemporary Urban Development. International Journal of Urban and Regional Research 24(1), pp.183-200.

Granovetter, M. (1973) "The Strength of Weak Ties," American Journal of Sociology 78, pp.1360-1380.

Grubesic, T., & O'Kelly, M. (2002) Using Points of Presence to Measure Accessibility to the Commercial Internet. The Professional Geographer 54.

Grubesic, T. H., O'Kelly, M. E., & Murray, A. T. (2002) "A Geographic Perspective on Commercial Internet Survivability." Telematics and Informatics 20(1): pp. 51-69.

Haggett, P., Cliff, A., & Frey, A. (1977) A Locational Analysis in Human Geography I: Locational Models. New York: John Wiley and Sons.

Hepworth, M. E. (1990) Geography of the Information Economy. London: Guilford Press.

Hepworth, M. E. (1991) Information technology and the global restructuring of capital markets, in S. Brunn, and T. Leinbach (Eds.), Collapsing Space and Time: Geographic Aspects of Communications and Information, pp. 132-148 (Chapter 12). London: HarperCollins

Holzmann, G. J., & Pehrson, B. (1994) The Early History of Data Networks. IEEE Computer Society Press. Los Alamitos, CA.

Hoover, E. M. (1974) An Introduction to Regional Economics University of Pittsburgh. New York. Alfred A. Knopf.

Hudsal, J., Sandvik, A., & Klingsheim, A. (1998). PloGIS-et verktoy I forvaltning av bilpark og planlegging av kjorerute (Norwegian, English: PloGIS-tool in car pool management and route planning, Coursework submitted in fulfillment of the one-year undergraduate study in GIS, Telemark College, Norway.

Hugill, P. J. (1999) Global Communications Since 1844: Geopolitics and Technology. Baltimore: Johns Hopkins University Press.

Hurley, D., & Keller, J. H. (1999) The First 100 Feet: Options for Internet and Broadband Access. Cambridge, MA. MIT Press.

Huston, G. (1999) Interconnection, Peering and Settlements: ISP Survival Guide. New York: John Wiley & Sons.

International Engineering Consortium [IEC] (2002) Optical Networks. http://www.iec.org/online/tutorials/opt_net/topic01.html January 28, 2003.

Jeong, H., Tombor, B., Albert, R. N. Oltvai, Z. N., & Lbarabasi, A. (2000) The Large-Scale Organization of Metabolic Networks. Nature 407, 651.

Kansky, K. (1963) "Structure of Transportation Networks," University of Chicago, Department of Geography Research Paper No. 84.

Kellerman, A. (1993) Telecommunications and Geography. London: Belhaven.

Kellerman, A. (2002) The Internet on Earth: A Geography of Information. John Wiley & Sons, New York, NY.

Kenny, M , & Curry, J. (2001) Beyond Transaction Costs: E-Commerce and the Power of the Internet Dadaspace, in T. Leinbach & S. Brunn (Eds.), Worlds of E-Commerce (Chapter 3). West Sussex, England: John Wiley & Sons.

Kochen, M. (1989) The Small World. Ablex Publishing Corporation, Norwood New Jersey

Kotkin, J. (2000) The New Geography: How the Digital Revolution is Reshaping the American Landscape. Random House, New York, NY.

Kotval, Z. (1999) Telecommunications: A Realistic Strategy for the Revitalization of American Cities. Cities: 16(1), pp.33-41.

Krebs, V. (2004) How to Do Social Network Analysis. http://www.orgnet.com/sna html.

Lalanne, L. 1863: Essai d'une theorie des reseaux de chemin de fer, fondee sure l'obeservation des taits et sur les lois primordiales qui president au groupement des populations. Comptes Rendus Hebdomadaires des Seances de l'Academie des Sciences, 42, pp.206-210.(3.1).

Langdale, J. (1989) The Geography of International Business Telecommunications: The Role of Leased Networks. Annals of the Association of American Geographers 79, pp.501-522.

Langdale, J. (2001) Global Electronic Spaces. Singapore's Role in the Foreign Exchange Market in the Asia-Pacific Region. Worlds of E-Commerce. eds. Leinbach & Brunn. John Wiley & Sons, Baffins Lane, Chichester West Sussex, England (Chapter 11).

Lawrence, S., & Giles, C. L. (1998), Searching the World Wide Web. Science 280, p.98.

Lawrence, S., & Giles, C. L. (1999), Accessibility and Distribution of Information on the Web. Nature 400, p.107.

LeBow, I. (1995) Information Highways and Byways: From the Telegraph to the 21st Century. New York: The Institute of Electrical and Electronics Engineers.

Leinbach, T. R., & S. D. Brunn (2001) Worlds of E-Commerce: Economic, Geographical and Social Dimensions. New York: John Wiley & Sons.

Liljeros, F., Edling, C. R., Amaral, L. A. N., Stanley, H E., & Aberg, Y. (2001) The Web of Human Sexual Contacts. Nature 411, 907.

Longcore, T. R., & Rees, P. W. (1996) Information Technology and Downtown Restructuring: The Case of New York City's Financial District, Urban Geography. 17(4), pp. 354-372.

MacDonald, A. & ESRI (2002) Building a Geodatabase. GIS by ESRI. Redlands, California

Malczewski, J. (1999) GIS and Multicriteria Decision Analysis. John Wiley & Sons, New York.

Malecki, E. J. (2000) The Internet: Its Economic Geography and Policy Implications. Paper prepared for Global Economic Geography Conference, Singapore, December.

Malecki, E. J. (2002) The Internet Age: Not the End of Geography, in D. Felsenstein & M. J. Taylor (Eds.), Promoting Local Growth: Process, Practice and Policy. Aldershot: Ashgate.

Malecki, E. J., & Gorman, S.P. (2001) Maybe the Death of Distance, but Not the End of Geography: The Internet as a Network, in S. D. Brunn & T. R. Leinbach (Eds.), Wired Worlds of Electronic Commerce, pp. 113-134. London: John Wiley.

Malecki, E. J., & McIntee, A. J. (2000) Access Ramps to the Internet: The Geography of Telecommunications Gateways in the U.S., unpublished paper.

McIntee, A. (2001) Urban Interconnections and the Geography of Colocation Facilities. Unpublished Masters Thesis, University of Florida.

Milgram, S., 1967, The Small World Problem. Psychology Today 1, 60

Montoya, J. M., & Sole, R.V. 2000, Small world Patterns of Food Webs. Preprint cond-mat/0011195.

Moss, M. L. (1991). The Information City in the Global Economy, in J. Brotchie, M. Batty, P. Hall, & P. Newton (Eds.), Cities of the 21st Century: New Technologies and Spatial systems, pp. 181-191. Melbourne: Longman Cheshire.

Moss, M. L. (1998) Technology and Cities. Cityscape: A Journal of Policy Development and Research. 3(3), 107-127.

Moss, M. L., & Townsend, A. M. (1998) Spatial Analysis of the Internet in U.S. Cities and States. Paper given at-Technological Futures-Urban Future Conference at Durham, England, April 23-25. http://urban.nyu.edu/research/spatial-analysis/spatial-analysis.pdf

Moss, M. L., & Townsend, A. M. (2000) The Internet Backbone and the American Metropolis. The Information Society 16, pp.35-47.

National Aeronautics and Space Administration (NASA) Earth Observatory collection of unique images, courtesy of the U.S. Air Force 2003 (source:. Chris Elvidge, U.S. Air ForceImage Satellite & Sensor DMSP www nasa.gov. Acquired August 15, 2003.

National Research Council, Computer Science and Telecommunications Board (2001) The Internet's Coming of Age. Washington, DC: National Academy Press.

Newman, M. E. J. (2001a) The Structure of Scientific Collaboration Networks. Proc. Natl. Adad. Sci. U.S.A. 98, 404

Newman, M. E. J. (2001b) Scientific Collaboration Networks: I. Network Construction and Fundamental Results. Physical Review E 64, 016131

Newman, M. E. J. (2001c) Scientific Collaboration Networks: II. Shortest Paths, Weighted Networks, and Centrality. Physical Review E, 64 016132

Newman, M. E. J. (2003) The structure and Function of Complex Networks: Arxiv: cond-mat/0303516 v1 March 25 http://arxiv.org/PS_cache/cond-mat/pdf/0303/0303516.pdf

Newman, M. E. J., Strogatz, S. H., & Watts, D. J. (2001) Random Graphs with Arbitrary Degree Distributions and their Applications. Physical Review E 64, 026118.

Odlyzko, A. (2000) The Internet and Other Networks: Utilization Rates and Their Implications. Information Economics and Policy 12, pp.341-365.

O'Kelly, M., & Grubesic, T (2002) Backbone topology, access and the commercial Internet, 1997-2000 Environment and Planning B 29 (forthcoming)

Orlando Sentinel (2001) Wired for the Americas: Miami Hub a Gateway for the E-Age. Sunday June 17th, Section H, pp. 1 & 3.

O'Sullivan, D. & Unwin, D. J (2003) Geographic Information Analysis. New York: John Wiley & Sons.

Ozana, M. (2001) Incipient spanning cluster on small-world networks Europhysics Letters Sept 15, 2001. 55, (6), pp. 762-766.

Pastor-Satorras, R., Vazquez, A., & Vespignani, A. (2001) Dynamical and Correlation Properties of the Internet. Preprint cond—mat/0105161.

Perrucci, R., & Potter, H. (1989) Networks of Power: Organizational Actors at the National, Corporate, and Community Levels Hawthrone, NY: Aldine de Gruyter.

Redner, S., (1998) Structural Properties of Scale-Free Networks. Eur. Phys. J. B 4, 131

Renesey 2003 SPECIAL REPORT. Blackout Results in Widespread Network Outages http://www.renesys.com/news/2003-08-14/blackout html

Rosenberg, N. (1994) Exploring the Black Box: Technology, Economics, and History. Cambridge: Cambridge University Press.

Rosenblum, S. (2003, August 15) For a Suddenly Powerless Manhattan, a 'Snow Day in August'. New York Times.

Sassen, S. (1991) The Global City: New York, London, Tokyo. Princeton NJ: Princeton University Press. (Chapters 8, 13)

Sassen, S. (1994) Cities in a World Economy. London: Sage; Thousand Oaks, CA: Pine Forge Press. (Chapters 5, 11)

Sassen, S. (1995) The state and the global city: notes towards a conception of place-centered governance. Competition and Change, 1(1), 31-50. (Chapter 8)

Sassen, S. (1996) Cities in the global economy. Proceedings of the Seoul Metropolitan Fora 96, Seoul. (Chapter 10).

Sassen, S. (1999) Global financial centers. Foreign Affairs, 78 (1), pp.75-87. (Chapter 11).

Sassen, S. (2000) Cities in a World Economy. Thousand Oaks, CA: Pine Forge Press.

Semple, K. (2003, August 14) Bush Says There is No Sign of Terrorism. New York Times.

Schiller, D. (1999) Digital Capitalism: Networking the Global Market System. Cambridge, MA: The MIT Press.

Shimbel, A. (1953). Structural Parameters of Communication Networks. Bulletin of Mathematical Biophysics, 15, 501-507.

Standage, T. (1998) The Victorian Internet. New York: Walker.

Steyvers, M., & Tenenbaum, J. B. (2001) preprint, www-psych.Stanford.edu/~jbt/

TeleGeography (2000) Resources: The Internet Exchange Points Directory. Washington, DC: TeleGeography, Inc. www.telegeography.com/ix/northamerica.html

TeleGeography (2001) Global Telecommunications Traffic Statistics & Commentary. Washington, DC: TeleGeography, Inc.

Terremark Worldwide (2000, May 22) Getting Heat in Miami. www costargroup.com/body/stories/2000/May/Telecomhotel asp

Thomas, G., & Wyatt, S. (1999) Shaping Cyperspace—Interpreting and Transforming the Internet. Research Policy 28, 681-698.

Townsend, A. (2001) The Internet and the Rise of the New Network Cities, 1969-1999. Environment and Planning B: Planning and Design pp.28, 39-58.

Vazquez, A. (2001) Statistics of Citation Networks. Preprint cond-mat/0105031

Von Geenhuizen, M., & Nijkamp, P. (2001) Electronic Banking and the City System in the Netherlands, in T. R. Leinbach & S. D. Brunn (Eds.), Worlds of E-Commerce John Wiley & Sons, Baffins Lane, Chichester West Sussex, England (Chapter 10).

Wagner, A., & Fell, D. (2000) Technical Report No. 00-07-041, Santa Fe Institute.

Warf, B., & Purcell, D. (2001) The Currency of Currency: Speed, Sovereignty, and Electronic Finance, in T. R. Leinbach & S. D. Brunn (Eds.), Worlds of E-Commerce (Chapter 12). West Sussex, England: John Wiley & Sons.

Wasserman, S., & Faust, K. (1994) Social Network Analysis. Methods and Applications. Cambridge: Cambridge University Press.

Watts, D. J (2003) Six Degrees: The Science of a Connected Age. New York: W. W. Norton & Company.

Watts, D. J., & S. H. Strogatz, 1998, Collective dynamics of small-world networks. Nature, 393, 440-442.

Wheeler, D. C., & O'Kelly, M. E. (1999) Network Topology and City Accessibility of the Commercial Internet. The Professional Geographer 51, 327-329.

Wheeler, J. O., Aoyama, Y., & Warf, B. (Eds.) (2000) Cities in the Telecommunications Age. The Fracturing of Geographies. London: Routledge

Williams, R. J., Martinez, N. D., Berlow, E. L., Dunne, J. A., & Barabasi, A-L. (2000) Santa Fe Institute Working Paper 01-07-036.

Wilson, M. (2001) Location, Location, Location: The Geography of the Dot Com Problem. Environment and Planning B: Planning and Design 2001, Volume 28, pp. 59-71.

Wilson, M. (2002) Chips, Bits, and the Law: An Economic Georgraphy of Internet Gamblin. Paper presented at the Annual Meeting of the Association of American Geographers (AAG), Los Angeles.

Wilson, M. I., & Corey, K. E. (Eds.) (2000) Information Tectonics. Chichester, UK: John Wiley & Sons.

Yook, S., Jeong, H., & Barabasi, A-L. (2001a) Modeling the Internet's Large-Scale Topology Reprint cond-mat/0107417

Yook, S., Jeong, H., & Barabasi, A-L. (2001b) Untitled. Unpublished.

Zeiler, M (1999) Modeling Our World The ESRI Guide to Geodatabase Design Redlands, CA: Environmental Systems Research Institute, Inc.

Zook, M A. (2000) The Web of Production: The Economic Geography of Commercial Internet Content Production in the United States. Environment and Planning A, 32, pp.411-426

Zook, M. (2001) Connected Is a Matter of Geography. Networker 5, 13-17.

BIOGRAPHICAL SKETCH

Angela McIntee was born to Paul and JoAnn McIntee in October 1978. She is the middle child of three, with two brothers. Andrew McIntee received his master's degree from the University of Florida in August 2001. Marc McIntee is currently working toward his bachelor's degree, also at the University of Florida.

Angela's interests lie in the telecommunications field, specifically Internet infrastructure and wireless communication.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Timothy J. Fik, Chair
Associate Professor of Geography

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Michael W. Binford
Professor of Geography

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Paul D. Zwick
Associate Professor of Urban and Regional Planning

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Ilir Bejleri
Assistant Professor of Urban and Regional Planning

This dissertation was submitted to the Graduate Faculty of the Department of Geography in the College of Liberal Arts and Sciences and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

May 2004

Dean, Graduate School